Revealing attention - how eye movements predict brand choice and moment of choice

Martinovici, A.

DOI:
10.26116/center-lis-1937

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Revealing Attention

How Eye Movements Predict
Brand Choice and Moment of Choice

Proefschrift

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. K. Sijtsma, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op dinsdag 17 december 2019 om 13.30 uur door

Ana Martinovici

geboren te Brăila, Roemenië

Promotor:

Prof. dr. F.G.M. Pieters

Copromotor:

Dr. R.J.A. van der Lans

Promotiecommissie:

Prof. dr. ir. B.J.J.A.M. Bronnenberg

Dr. H. Datta

Prof. dr. E. Gijsbrechts

Dr. A. Lemmens

Prof. dr. S. Misra

Prof. dr. ir. A. Smidts

Acknowledgements

This thesis is about the role of attention in consumer choice. Before discussing that, I would like to draw your attention to the people who made this thesis possible.

One can be a PhD student only if one has a PhD advisor. Hence, I start by expressing my gratitude towards Rik. Thank you. For everything that I've learned from you in the past years, for your passion and dedication to research. When we met, I wasn't even considering doing a PhD, as my plan at the time was to go back to industry and get a 'normal' job. Silly plan! You opened the door to academic research and encouraged me to step over the threshold. For this, I will forever be grateful. My life would be very different without attention, decision making, Bayes, and R. And I've got to discover all of these and much more only because of you.

Ralf, I am grateful to have you as a co-promoter. Your suggestions have improved the thesis and helped me develop skills that will also benefit future projects. Thank you for your support throughout the intense final stages of the thesis writing process.

I am grateful to Bart Bronnenberg, Hannes Datta, Els Gijsbrechts, Aurelie Lemmens, Sanjog Misra, and Ale Smidts for being part of the PhD committee. You've provided me with many useful suggestions and have contributed to the current version of this dissertation. I am truly fortunate to have such a wonderful committee.

People say it takes a village… or more specifically, it takes a department to grow a PhD student into a researcher. For that, I want to thank the members of the Marketing department at Tilburg University. You provided so many valuable learning moments and I've tried to learn as much as possible from your feedback during the PhD camps and razor-sharp questions during talks. The training offered during the RM and PhD program equips those who successfully leave the nest with much-needed research skills. For all of this, the Tilburg group will always hold a special place in my heart.

I finished my dissertation in the Marketing Management department at RSM, Erasmus University. I want to say a big 'Thank you' to all my colleagues at RSM. You've made it possible for me to finish the last part of the thesis by offering me the best mix of support and encouragement one could wish for. I am happy to be part of the team and look forward to what's to come.

During my RM and at the start of my PhD I was lucky to become friends with an awesome group of Nutella fans. Max, Esther, Kris, thank you for all the moments we've shared together. Although we didn't have the best influence on each other's diets, I like to think the bike rides we took compensated for all the vending machine snacks. Your support and encouragement have helped me get back in the saddle and that made a big difference. On that note, many thanks to Agnes and Milena as well.

Esther, we got to know each other due to our studies, but throughout the years we've become more than colleagues and grew into academic sisters. Thank you for being a friend, for inviting me into your family, for all the Christmas meals we've shared together, for learning how to count in Dutch over a game of Bingo with your grandmother, parents, and family. Your family has adopted the odd Dutch-not-speaking guest and made me feel welcome. Thank you and them for all this. See you soon!

Liz, we’re probably the best example that conscious uncoupling is real. Thank you for making this work and providing stability in a time when I very much needed it. Iulia, Alina, I am happy that even after all these years the Plevnei group keeps going. Looking forward to set the tone on the dance floor next year. Many thanks go to all the people who’ve helped me, sometimes unknowingly, get through a rough patch: Jennifer Argo, George Knox and Otilia Boldea, Andriana and Ion, Abbelia, Diana, and Trang.

Mama, tata – thank you for encouraging me to invest time and effort in my studies. Vă sunt recunoscătoare pentru sacrifiicile pe care le-ați făcut de-a lungul anilor.

Finally, to Andreas: dankeschön für alles. I save the dankjewel for when I am able to use the Dr. title.

October 2019

Contents

Chapter 1

Introduction

1.1 Motivation

Remember the last time you made an online purchase without having previously seen the real physical product. Consumers frequently do this, for example, when they order food from a new restaurant or purchase a device that is not available in physical stores. If you are in academia, then it is probably easier to remember booking a room in a hotel you have not visited before, as is usually the case when going to a conference. This should be an easy choice, even more so if recommended conference hotels are available. All you need to do is search for information about the available options and then select the one that you think is best. Performing this search online gives you easy access to large amounts of information about the alternatives – you can check not only prices and hotel amenities, but also pictures, ratings, and reviews. However, acquiring all the available information about the alternatives involves time you probably prefer to spend in a different way. Therefore, you focus your attention primarily on information relevant for your choice at the time.

What information is relevant differs both within the same consumer over time and between consumers at the same point in time, as it depends on their specific needs and goals. For example, PhD students pay more attention to the price of the hotel room and the possibility of sharing it with a colleague, as this is relevant given their budget constraints. Those who need to present early in the morning are more likely to focus on distance to the conference venue, while others might find information about airport connections and hotel amenities most relevant.

These examples show that different consumers pay attention to different information, even when making a choice from the same set. More importantly, the information that consumers pay attention to reflects what influences their choice, even before they click on "book this room" or "buy this brand". Such differences in attention to information about the alternatives can be observed in many domains and are in no way specific only to academics booking a hotel room for a conference. As consumers aim to make a purchase, they search for those alternatives that are aligned with their preferences and inspect the information that is most relevant for their goal. This means that the search for information that precedes consumer choice is closely linked to what consumers find useful, important, relevant. Companies such as Google, Amazon, and Facebook, integrate this in their machine learning algorithms that personalize search results, offer product recommendations, or deliver targeted ads, with the final goal of influencing the choices that their users make.

Consumer choice models and theories rest on various assumptions about information search and brand choice. First, information search prior to brand choice indicates what the consumer knows about the brands. Second, consumer preferences are revealed only by the choice of brand, not by the sequence of (micro) choices regarding what information to examine from moment-to-moment. This implies that it is impossible to predict what brand a consumer is going to choose in the absence of prior preference measurements (e.g. previous choices, preference ratings). As a result, current search-and-choice models can only offer a "choice post-mortem" on why the consumer chose a certain brand. This limits our understanding of how consumers make a brand choice and of the role that sequential information acquisition plays in decision making. This dissertation addresses this challenge and provides a model that infers consumer and brand-specific utilities from how the consumer inspects the alternatives during a single choice. This differs from standard choice models that specify brand utility as a weighted sum of product attributes and that need multiple observed choices in order to measure heterogeneous consumer preferences.

1.2 Eye Movements and Attention

The three empirical essays in this dissertation all use eye movements as indicators of information acquisition and attention processes (Pieters and Wedel 2007). When consumers inspect brands on a display, they make eye movements called saccades. During saccades, the gaze is rapidly redirected (20-50 msec.) between different locations on the display, while vision is actively suppressed to prevent blurring (Hutton 2008). Between saccades, the eyes are relatively still (for about 200-400 msec.) and focused on a specific location in space. These brief moments between saccades are called fixations and it is during these moments that the consumer acquires, by reading, the information presented on the corresponding area of the display (Rayner 1998).

Eye-movement data provide spatiotemporal information: (1) what areas of the display are inspected, (2) the moments when this takes place, and (3) the duration of these moments. This has three implications. First, eye-fixations during a task indicate which areas of the display the consumer attends to. Second, the sequence of eye-fixations and saccades reflects the moments when consumers are interested in the specific information presented at different locations on the display. While consumers aim to inspect relevant information, areas that contain salient stimuli can attract their attention (van der Lans, Pieters, and Wedel 2008a). If such salient stimuli do not provide task-relevant information, consumers do not inspect them further and move their eyes to other areas. This relates to the third implication. Namely, that the amount and pattern of eye movements reflect what information in areas of the display is relevant for what the consumer aims to achieve during the task. In the previous conference hotel example, a consumer who aims to choose an alternative that is well connected to the airport inspects information about travel time and airport-shuttle services more than hotel amenities (i.e., what is inspected). While stimuli such as pictures and review ratings displayed as colored stars are likely to be fixated early on (i.e., the moment when it takes place), they are not directly relevant to the consumer and receive a lower number of eye-fixations (i.e. duration of attention) than other, more relevant, areas.

The link between eye movements and goal-directed, or top-down, attention is supported by studies in scene viewing, advertising (Pieters and Wedel 2007; Wedel, Pieters, and Liechty 2008), and search (van der Lans, Pieters, and Wedel 2008b). For example, participants instructed to memorize ads allocate more attention to the body text, pictorial, and brand design objects than participants instructed to explore the ads freely, while those following a brand-learning goal attend more to the body text but reduce their attention to pictorial design (Pieters and Wedel 2007). The effect of processing goals on attention is manifest even when the design of the display could lead to more homogenous attention patterns between consumers who inspect it. The results of an eye-tracking study that decomposed the effects of brand salience on search show that about two thirds of brand attention is goal-directed while only one third is stimulus-driven (van der Lans et al. 2008b).

To summarize, previous studies show that eye movements are closely linked to the goal that participants have. Therefore, eye movements reflect how consumers divide their goal-directed attention between the brands on display. During choice, consumers move their eyes from moment to moment in order to inspect the information displayed in different locations. This implies that the sequence of eye movements reflects three important aspects of attention: selection (what was attended), pattern (at what moments), and duration (for how long). Let y_{ibt} be the number of eye movements[1] of consumer i on the display area corresponding to brand b[2] at moment t, where $t = 1$ at the start of the choice task and $t = T_i$ at the end of the task, when brand choice is expressed. We use y_{ib}^{t} to indicate the sequence

[1] The type of eye-movements (e.g. fixations, saccades) used in each of the empirical essays is defined in the respective chapter.

[2] The level at which eye movements and attention are modelled can easily be changed (e.g. in Chapter 4 we use eye movements on attribute information and on brand-and-attribute cells).

$(y_{ib1}, \ldots, y_{ibt})$ of eye movements of consumer i on the display area corresponding to brand b from the start of the task and until moment t. We formalize the link between eye movements (y) and attention (θ) in equation 1. More specifically, the sequence of eye movements of consumer i on the display area corresponding to brand b up to moment t reflects the attention allocated to the brand (θ_{ibt}):

$$y_{ib}^{t} = \theta_{ibt} + \xi_{ibt}. \tag{1}$$

Brand and consumer characteristics that are expected to influence attention can be included in the model (X_{ib0} and X_{i00}, respectively) and their effects are captured by η_{01t} and η_{10t}. For example, brands that the consumer is already familiar with (e.g. based on prior ownership) are expected to attract more attention initially as they are easier to recognize than yet unknown brands. In line with prior research, consumer characteristics such as decision goals are expected to influence attention. Then, consumer and brand specific attention is:

$$\theta_{ibt} = \eta_{01t}X_{ib0} + \eta_{10t}X_{i00} + \epsilon_{ibt}. \tag{2}$$

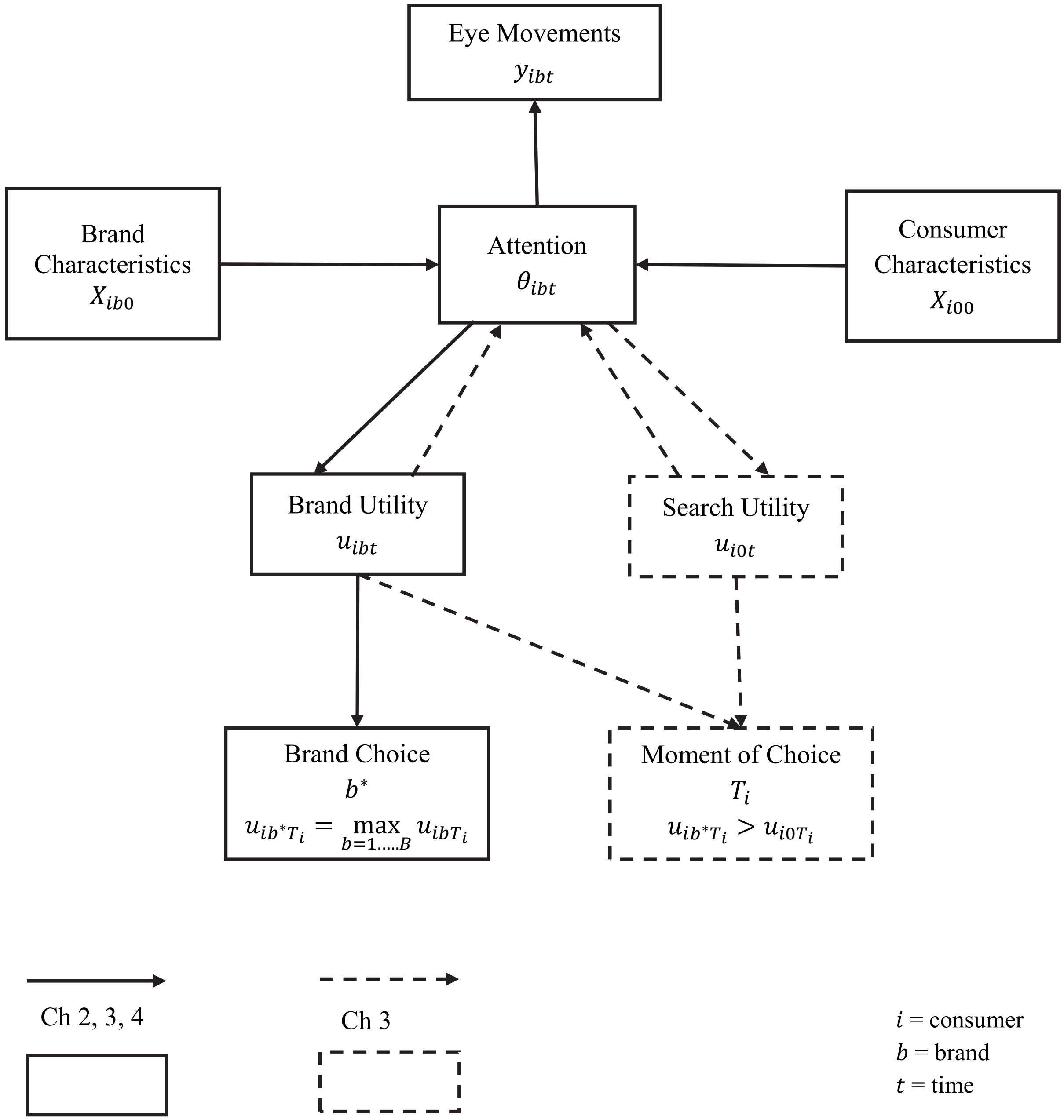
The three essays in this dissertation specify that eye movements are not perfect indicators of attention by accounting for unobserved sources of heterogeneity (ϵ_{ibt}, e.g. consumers relying on different types of information, such as textual or pictorial (van der Lans et al. 2008a)) and measurement error (ξ_{ibt}, discussed in detail in sections 2.2.1 and 3.2.1).

Equations 1 and 2 formalize the link between eye movements and attention, the first component of our model. Figure 1.1 offers a visual representation of the model and summarizes which links are tested in the three empirical essays (Chapters 3, 4, and 5).

1.3 Attention, Brand Choice, and Moment of Choice

Standard choice models assume that consumers use all the available information to calculate the utility of each brand and then select the best alternative in the set. However, consumers rarely inspect all the information available to them, even when this is presented at the same

Figure 1.1 Framework of this thesis: eye movements, attention, utility, and choice



Note: Examples of brand characteristics: prior brand ownership (Chapter 2); position on display (Chapters 2 and 3); and attribute levels (Chapter 4). Examples of consumer characteristics: smartphone ownership (Chapter 2); between subjects manipulated information complexity (Chapter 2), decision goals (Chapters 3 and 4), and time pressure (Chapter 4).

To avoid repetition, the following terms are used interchangeably in this dissertation: (1) brand, (choice) alternative, and (choice) option; and (2) attributes, (brand) characteristics, (brand) features. One brand corresponds to one alternative presented on the screen. If two or more alternatives have the same brand name, but different other attributes (e.g. iPhone XR and iPhone XS Max), the model considers them as different brands with a shared brand characteristic (brand name).

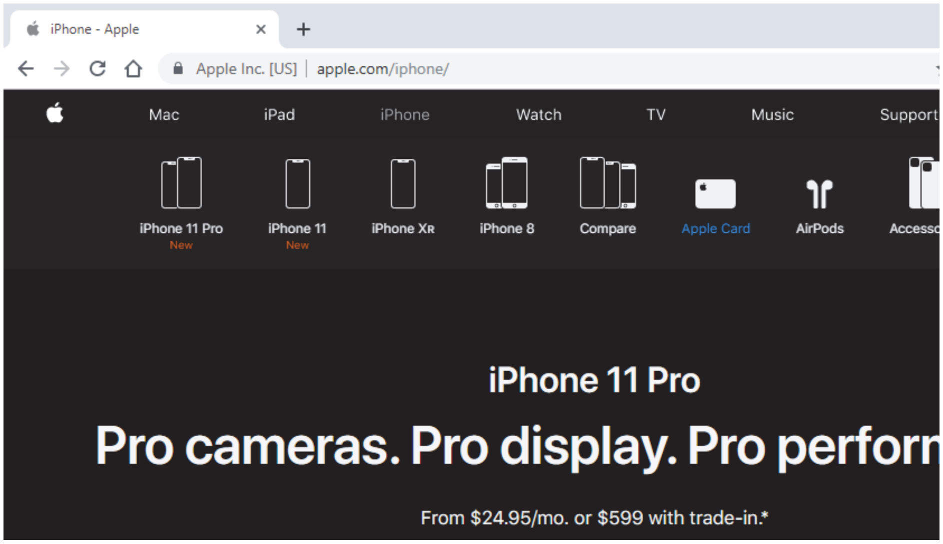
time on a display (e.g. attribute by brand matrices commonly used by comparison websites and online retailers). This is supported by observational data that covers weeks of online search (Bronnenberg, Kim, and Mela 2016) and controlled studies that facilitate information acquisition by presenting brands side-by-side (Meißner, Musalem, and Huber 2016; Shi, Wedel, and Pieters 2013).

Before consumers visually inspect the display, they are uncertain about the specific brands and attributes that they are going to see, even though they might have an expectation about them. For example, consumers who click 'Compare' while browsing the Apple store for an iPhone (Figure 1.2) can expect to see information about attributes such as camera, battery, and price for a limited number of models. However, they are uncertain about the exact location where this information is presented and the specific attribute levels that correspond to each of the models. To the extent that consumers find an attribute important or are interested in one of the models, they benefit from reducing uncertainty about it. Then, they use their eyes to find the area where that information is displayed and inspect it. Section 1.3.1 describes four types of uncertainty that consumers experience during choice, and how these are accounted for by previous research where the role of attention is implied.

What consumers choose to inspect is aligned with what they find important and what impacts their decision. Failing to account for what information consumers inspect during choice leads to biased measurements of their preferences and incorrect inferences about their decision process. This has motivated the development of new theories and models, such as sequential sampling models (SSM) and rational inattention theory (RIT) on which the essays in this dissertation build on (SSM and RIT described in 1.3.2 and 1.3.3). In section 1.3.4 we describe how the models in this dissertation specify the link between attention and utility.

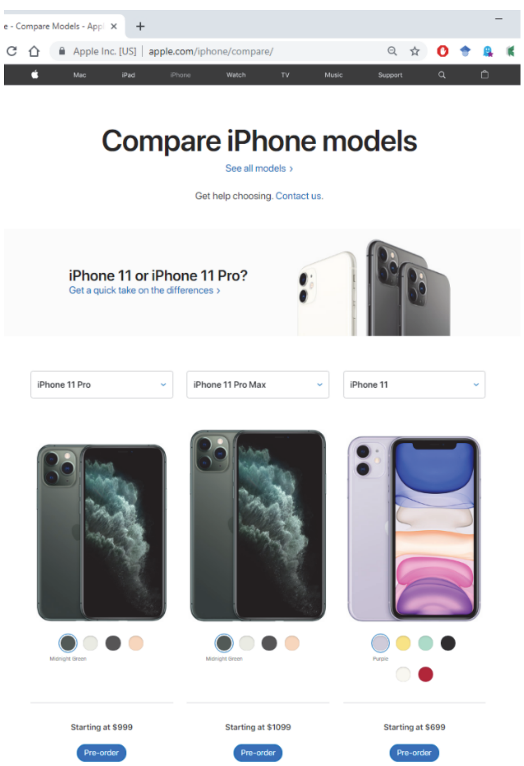
Figure 1.2 "Compare" section at apple.com/iphone

Start page

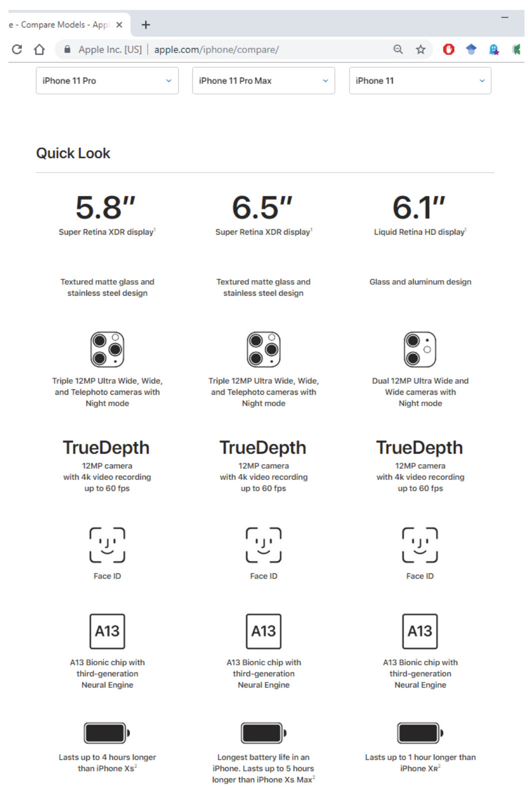


Information available after clinking "Compare"

First screen



After scrolling down the page



Note: Screenshots taken on 18 September 2019.

1.3.1 Uncertainty during Brand Choice

When choosing one brand from a set of several multiattribute alternatives, consumers can be uncertain about: (1) the description of a brand on a specific attribute, (2) the importance of an attribute, (3) the overall utility of a brand, and (4) which of the brands has the highest utility in the set. For a consumer who wants to purchase a bike, the four types correspond to the following questions: (1) what is the material of the frame (e.g. aluminum, carbon fiber, steel) or the model/type of the derailleur[3], (2) how important is the material of the frame, or the derailleur, (3) what is the total utility that the consumer would derive from bike A, and (4) which of the bikes in the set provides the best utility. While type I can be resolved by reading the description of the bikes, this might not be enough for types 2-4. In order to understand the importance of an attribute (type II), the consumer needs to understand what benefits it provides (e.g. maintenance, performance). Benefits are usually not explained next to attributes and often require specific knowledge about the product category. Even for bike enthusiasts who have information about all the bike characteristics and know their importance, a test ride is needed to determine if the bike size and configuration is a good fit (type III uncertainty). Hence, resolving type II-IV uncertainties might require prior experience with the category, learning through consumption or test-driving a product before purchase, in addition to careful inspection of the information on display.

Type I: Uncertainty about the description of a brand on a specific attribute. Before acquiring any information about a brand, consumers are uncertain about its attributes. After inspecting some of the attributes, consumers can form expectations for the remaining ones, to the extent that these are correlated. For example, after discovering that a digital camera has exceptionally good video recording performance, a consumer can expect a higher price. However, the consumer can only be certain of this after inspecting the price of that digital

---

[3] Derailleurs are used to change the gears on the cassette, together with the shifters.

camera. In this example the consumer can easily reduce uncertainty by reading the information on display. However, there are situations when this is more difficult to achieve. For example, when attribute information is in an unknown language (e.g. tourists shopping in a foreign country) or when buying in a completely new category (e.g. first-time parents shopping for baby products).

Information that consumers do not know cannot influence their evaluation of a brand. For example, consumers consider the contribution of the CPU to the utility of a laptop only if they know that information. Hence, accounting for what brand information consumers use to evaluate the brands improves preference measurements (Meißner et al. 2016; Yang, Toubia, and de Jong 2015).

Type II: Uncertainty about the importance of an attribute. Even though consumers are certain about the description of a brand on a specific attribute, they can be uncertain about its contribution to the brand's utility. For example, after learning that brand A has a battery life of 36 hours, consumers can be uncertain about the benefit of this level of battery life. Consumption benefits are more difficult to evaluate for consumers who make a purchase in a new category or for an experience good (Bronnenberg and Dubé 2017). When this type of uncertainty is present, consumers use other sources of information to evaluate the utility of the choice alternatives, such as advertising (Erdem, Keane, and Sun 2008; Kotowitz and Mathewson 1979). Hence, companies can help consumers resolve or even prevent this type of uncertainty by providing cues (e.g. banner ads) that help consumers to remember previously seen advertisements for products in the respective category.

Type III: Uncertainty about the overall utility of a brand. Even when consumers have all the brand information and know the subjective value that they derive from each of the attributes, they can be uncertain about its overall utility. For example, if the brand is described by multiple attributes and consumers need to consolidate the attribute partworths or if there are interactions between attributes or some of the attributes are difficult to trade off (Bettman, Luce, and Payne 1998).

Models of sequential search and choice, as well as learning models, account for this type of uncertainty. Search and choice models assume that consumers have an expectation of the utility provided by the attributes of the brand prior to search but are uncertain about the realized utility (Kim, Albuquerque, and Bronnenberg 2010). During search, consumers resolve this uncertainty one brand at a time. Learning models account for consumers having incomplete information about attributes that they discover over time. Because learning models usually account for product quality as an overall attribute, we discuss them in relation to type III uncertainty and not the previous two. An important characteristic of this class of models is that consumers learn about the quality of one brand, but this knowledge does not directly influence the utility of the other brands. In contrast, if consumers reduce type II uncertainty (e.g. a consumer updates price sensitivity), then all the brands are impacted. In a review of the literature, Ching, Erdem, and Keane (2013) discuss four key dimension that characterize different types of learning models. One of these dimensions is the source of information consumers use to learn about attributes over time. For example, consumers can learn by consuming the product (i.e. after purchase), but they can also learn from exogenous signals of quality such as advertising. Similar to RIT, learning models assume that the reduction in consumer uncertainty is independent of the order or the timing of these exogenous signals.

Type IV: Uncertainty about which brand has the highest utility in the set. The previous three types of uncertainty are at the level of an attribute and/or brand, while type IV is at the choice set level and is related to the difference in utility between the brands. Two or more brands can have similar levels of utility because (1) their attributes are similar or (2) they outperform each other on attributes that are difficult to trade off. Examples, respectively:

(1) choosing the colour of a smartphone or laptop – while essentially unimportant and irrelevant, the consumer needs to resolve this uncertainty, and (2) one brand is expensive and offers high quality, while the other is cheaper but has lower quality. Consumers faced with such a difficult trade off can switch towards a more simple decision strategy (e.g. lexicographic rule, satisficing) (Luce, Payne, and Bettman 2000) and thus eliminate the difficulty of choosing the best brand in the set, or can put more effort into differentiating the brand utilities.

1.3.2 Sequential Sampling Models (SSM)

SSMs have been used primarily to understand processes that take place during perceptual decisions (Ratcliff 1978), such as whether a visual stimulus displays a square or not, or whether it displays a chair or a table. Recent developments, such as the attentional drift-diffusion model (aDDM) (Krajbich, Armel, and Rangel 2010; Krajbich and Rangel 2011) and multialternative decision field theory (MDFT) (Roe, Busemeyer, and Townsend 2001), have generalized the model to value-based choice. The core assumption of these models is that participants' sequential sampling of information influences how evidence in favor of the choice alternatives accumulates until a threshold is reached and choice is expressed (Ratcliff and Smith 2004). This implies that alternatives that are fixated on more frequently are more likely to be chosen.

While there are several classes of SSMs, of which the aDDM and MDFT are just two examples, these models share the following three components: drift rate, decision threshold, and starting point. The drift rate is the average amount of evidence accumulated per unit of time. The decision threshold is the amount of evidence that needs to be accumulated in order for choice to be expressed. The starting point captures prior preferences for one of the choice alternatives. The aDDM specifies that brands accumulate evidence at a higher rate when they are fixated on as compared to when the consumer fixates on a competing brand (Krajbich et al. 2010). This implies that brands accumulate evidence even when they are not looked at, albeit at a lower rate. aDDM applications use a consumer-invariant decision threshold that is fixed prior to the start of the decision process and remains constant throughout the task. Because these models require measurements of preference ratings for all of the items that participants choose between, it is straightforward to adjust the starting point accordingly. However, this is not needed in practice as participants are asked to make choices between two or at most three simple items (e.g. chocolate bars, snacks) with similar preference ratings (Krajbich et al. 2010; Krajbich and Rangel 2011). The elementary nature of these choices makes it possible to ask participants to do hundreds of choice tasks, which are needed to fit the models. While such applications offer valuable insights into the “computational and psychological processes that guide simple choices” (Krajbich et al. 2010, p. 1296), it is not immediately obvious how they can be extended to complex choices between multi-attribute brands for which no prior preference measurements are available.

Applications of the aDDM focused on choice between simple alternatives (e.g. snacks) that participants know and like (Krajbich et al. 2010; Krajbich and Rangel 2011). Because participants choose between alternatives that they are already familiar with, the tasks are more similar to perceptual decision making or brand search than they are to brand choice. Specifically, they only need to identify what brands are on display and choose the one that they assigned a higher preference rating. Then, the two or three brands on display should have similar attention shares. We use the term *fair share* to refer to the share of attention that specific areas of the display (e.g. brands, attributes) are expected to receive if participants use eye movements only for information acquisition.

The aDDM literature uses the term choice bias to indicate that “controlling for value differences, the probability of choosing an item should increase with the excess time for which it is fixated” (p. 1294). Similarly, the term attention bias indicates that participants are more likely to fixate on items that they prefer and thus allocate a larger share of their attention to items that they eventually choose. To clarify, *bias* is used in the sense of inclination or interest in some items and does not imply that participants make an error in how they allocate attention or what brand they choose.

Decision field theory, originally developed for decision making under uncertainty (Busemeyer and Townsend 1993) has been extended to value-based choice (Roe et al. 2001). Similar to the aDDM, models in the multialternative decision field theory class (Roe et al. 2001) assume that choice alternatives accumulate evidence until a threshold is reached and choice is expressed, and that this accumulation process is modulated by attention. In addition, to be estimated, both theories require repeated choices per consumer (Berkowitsch, Scheibehenne, and Rieskamp 2014). While the aDDM is specified for unidimensional stimuli (Krajbich et al. 2010), such as chocolate bars or snacks, MDFT models focus on choice between alternatives described by multiple attributes, such as cars. Then, attention influences which attribute is under focus at the specific time. The model assumes that attention operates like a filter that selects the attribute on which the brands are compared from moment-to-moment. Importantly, at every moment during choice, after attention selects an attribute of interest, all the brands are inspected and compared on that dimension (Roe et al. 2001). So far, these models have primarily been used to simulate choice behavior and test whether MDFT is able to explain context effects (e.g. similarity, attraction, compromise).

1.3.3 Rational Inattention Theory (RIT)

RIT argues that because information acquisition is costly, it is rational for consumers to not pay attention to information that is less relevant for their choice. Then, consumers maximize utility given the limited information they have about the choice options. This leads to situations when “limits on attention impact choice” (Caplin and Dean 2015, p. 2183) and consumers choose a brand with lower utility than if they were to choose based on full information. Recent developments in this literature provide analytical models of optimal information-processing behavior (Matějka and McKay 2015; Steiner, Stewart, and Matějka 2017) that as far as we know have not been tested empirically for value-based choice.

Consumer behavior that is aligned with RIT satisfies two assumptions: no improving attention cycles (NIAC) and no improving action switches (NIAS) (Caplin and Dean 2015). The first assumption, no improving attention cycles, specifies that decision makers' attention allocation is rationalized by a cost function. The second assumption, no improving action switches, implies that consumers' choice of action is optimal given the information gathered. So far developments of these models have focused on analytical results that test if participants adjust their attention and actions in response to changes in incentives in line with predictions derived from the NIAC and NIAS assumptions. In order to be implemented, these tests require state-dependent stochastic choice (SDSC) data (Caplin 2016; Caplin and Dean 2015). SDSC data comprises of: (1) a set of actions that the consumer chooses from, (2) the utility of each of these actions in different states of the world, (3) the consumer's prior belief about the true state probabilities, and (4) the probability of observing different information signals given the true state of the world. A working paper on empirical tests of RIT offers a more concrete example of SDSC data (Dean and Neligh 2019). This paper uses an experiment in which participants are presented with 100 red and blue balls on the screen and then choose between two actions. The payoffs of the actions depend on the fraction of red balls on the screen (the true state of the world). Participants know the prior probability of the possible true states and can determine the state by counting the balls on the screen. Systematic manipulations of the true state of the world and of the payoffs should lead to changes in the strategies that participants use to choose an action. Changes that are in line with the NIAS and NIAC predictions support RIT.

1.3.4 Towards a Theory of Rational Attention (TRA)

Prior preference measurements at the consumer level are not easily accessible, especially for complex brands that are infrequently purchased. This creates a significant challenge in implementing the aDDM or RIT models for the type of brand choices that are common in marketing applications, as both these classes of models require prior measurements of consumer preference (the true state in RIT models corresponds to the utility match between the participants' underlying preferences and the characteristics of the brands in the set).

This dissertation offers a solution to this challenge, as we describe in this section. Our approach makes it possible to describe and quantify the role of eye movements in reflecting otherwise unobserved attention processes that are closely linked to the moment-to-moment accumulation of utility that takes place during brand choice. In neither of the three empirical essays do we (as researchers) a priori know the true underlying distribution of the state of the world (the utility match between each participant and brand on display). It is reasonable to assume that prior to any fixations on the display, participants do not know these utilities either, except when they would know which brands they are going to see and know them sufficiently well to have an estimate of their utilities. However, to the extent that consumer behavior is in line with RIT, both the allocation of attention between the brands and the resulting choice are utility maximizing. This implies that consumers postpone brand choice as long as they derive more utility from inspecting the brands than from selecting one of them as their final choice. Our model incorporates this by specifying two types of utilities (brand (u_{ibt}) and search (u_{i0t})) that consumer i compares at every moment $t = 1, \dots, T_i$.

If the consumer's choice of action is optimal given the information gathered (NIAC) and the attention allocation is rational (NIAS), then attention is closely linked to the utility that the consumer derives from moment-to-moment during the decision process. Building on

SSMs, we specify that the utility consumer i derives from choosing brand b at time t is a function of the attention allocated to that brand and an unobserved utility shock ε_{ibt}:

$$u_{ibt} = \beta * \theta_{ibt} + \varepsilon_{ibt}. \tag{3}$$

To summarize, the essays in this dissertation build on SSM, RIT, and research that documents the role of eye movements as indicators of attention. The models in the three essays of this dissertation share the following premises: (1) consumers allocate attention strategically, (2) consumers maximize utility given what they know about the brands, and (3) eye movements reflect these processes from moment-to-moment. Hence, consumers choose what to focus their attention on, rather than what to be inattentive to. While this could be described as rational attention, the models developed in this dissertation do not formally test rationality. The goal of this dissertation is to understand how eye movements reflect otherwise unobservable attention and utility accumulation processes that take place when consumers make a single brand choice from a set of complex alternatives. Whether consumers' moment-to-moment choices of what to fixate their eyes on are rational or not, our results show that the resulting allocation of attention *reveals* how consumers evaluate the brands and predicts choice. In chapter 5 we come back to this key idea and discuss how the models and results of this dissertation take us one step closer towards a Theory of Rational Attention, while acknowledging that many such steps remain.

In the next section we describe how each chapter focuses on specific model components. Then, each chapter presents in detail the developed model, the data on which it is calibrated, and the results.

1.4 Outline

The attention and brand choice model in Chapter 2 offers initial evidence to support the link between changes in attention over the course of the decision process and brand utility at the moment of choice. Importantly, it investigates which fundamental attention processes contribute to the accumulation of utility and brand choice. The model specifies the effects of information density on four types of consumer-and-brand specific attention, and tests competing mechanisms through which brand loyalty manifests itself in choice and attention. The results show that (1) certain types of attention (e.g. attention for integration) are better able to reflect brand utilities, (2) brand loyalty manifests itself via attention, and that (3) the link between attention and brand utility is stable across different information density displays. To keep the model tractable, we normalize decision duration by splitting it in four quarters and investigating changes in attention over these intervals.

Chapter 3 looks into the link between attention, brand choice, and moment of choice. The model specifies both brand and search utilities that change from moment to moment as more eye movements are observed. The model predicts both brand choice and the moment of choice, and the results show that the accuracy of these predictions is above chance even after only 30% of the decision time. This provides evidence that already early in the process eye movements predict both what will be chosen and when this choice is expressed. The chapter provides insights into consumer heterogeneity in decision thresholds and implicitly decision duration, and test different drivers of brand choice and moment of choice.

While in Chapters 2 and 3 eye movements reflect attention that is linked to overall brand utilities, Chapter 4 takes a different approach. The model used in this chapter decomposes brand utilities into two components that capture the importance of the attributes that describe the brands and the subjective value that the consumer attaches to the attribute levels corresponding to each of the brands on display. The results show that eye movements reflect not only how the consumer evaluates the brands, but also why some brands are preferred by identifying which attributes are considered more important over time.

There are six main novelties across the three empirical essays. First, the model in Chapter 3 predicts both brand choice and choice timing. This is possible because it explicitly models both brand and search utilities. Consumers compare these two types of utilities in order to determine when to express their choice and compare brand utilities in order to determine what brand to choose. When choosing a brand offers more utility than continuing to inspect the information on display, the consumer ends the decision process and expresses brand choice.

Second, predictions of brand choice and choice timing are made for new consumers. Different from SSM and RIT models, this approach does not require any prior preferences measurements (e.g. ratings or choices). Of course, if such prior measurements are already available, they can easily be included as prior information both when calibrating the model and when making predictions for new participants, as explained in section 3.4.4.

Third, the model can easily be adapted to extract brand utilities, attribute importance weights, and subjective attribute values, all consumer specific. These utility components are inferred from the allocation of attention between brands, attributes, or brand-and-attributes, which are reflected in eye movements.

Fourth, the model can accommodate choice sets of any size, unlike the aDDM which so far can be applied only to at most three alternatives. Because utility is a function of attention and not of attribute levels, the model does not suffer from the usual criticism of the independence of irrelevant alternatives property of logit models (Matějka and McKay 2015). When choice sets increase, consumers are more likely to ignore some brands, even more if they have very similar attribute levels. As a result, these brands receive little, if any, attention and as a result have low utilities and choice probabilities that do not impact the choice probabilities of the other brands in the set.

Fifth, the model predicts what and when consumers choose, before they implement the choice. It does so from moment-to-moment and updates the predictions as more data becomes available (described in section 3.4.4).

Sixth, the model is calibrated on a combination of eye movements and brand choice data, but predictions in chapters 3 and 4 are based on eye movements only. Currently, eye movements are not routinely collected by online platforms. However, recent developments that make it easier and more affordable to collect eye tracking data suggest that this could change in the future. For example, eye movements can be tracked using the camera of a laptop/PC/smartphone (Lopez et al. 2017) and many of these devices are fitted with infrared emitters and camera[4] that can improve tracking accuracy. Until such solutions are fully developed and adapted for eye tracking, other types of data (e.g. browsing) that reflect consumer interest and attention could be used. Chapter 5 discusses this in more detail as well as the ethical concerns that arise from being able to infer consumer preferences and attitudes.

[4] For example: iPhone X and later, iPad Pro models with A12X Bionic Chip (https://support.apple.com/en-us/HT208108); laptops that support facial recognition (Windows Hello); smartphones that use iris scans (Samsung Galaxy S8 / Note 8 or later).

Chapter 2

Eye Movements, Attention, and Utility Accumulation during Brand Choice

2.1 Introduction

This chapter investigates the idea that eye movements provide a unique window into fundamental but not directly observable processes of preference formation and utility accumulation during brand choice in information-rich environments. Because attention in such environments is a scarce resource, it enables us to test a number of predictions derived from rational inattention theory (RIT) (Maćkowiak, Matějka, and Wiederholt 2018; Matějka and McKay 2015).

Consumers make complex choices in information-rich environments, such as when choosing between different housing options, holiday destinations, household appliances, or smartphones. Even when all information is simultaneously available at a single location, such as a comparison website, consumers' limited attentional capacity prevents them from carefully devoting full attention to each of the choice options (Lohse and Johnson 1996; Shi et al. 2013; Willemsen, Böckenholt, and Johnson 2011). Early on, Simon (1971, pp. 40-41) expressed the challenge that consumers in an information-rich world face as follows: "… the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it." Based on this notion of the scarcity and costliness of attention, RIT posits that it is in the consumers' best interest to process information that they find useful and ignore or pay less attention to choice alternatives that seem less worth the effort. This implies that there should be a positive association between the attention that consumers devote to the alternatives in a choice set and their respective choice likelihoods. There is indeed evidence for such a positive association (Atalay, Bodur, and Rasolofoarison 2012; Chandon et al. 2009; Krajbich et al. 2012; Pieters and Warlop 1999). There is also evidence that consumers' attention during choice tasks, as measured by eye movements, reflects key cognitive processes that consumers engage in prior to expressing their choice (Al-Moteri et al. 2017; Arieli, Ben-Ami, and Rubinstein 2011; Glaholt and Reingold 2011; Lohse and Johnson 1996). Moreover, accounting for the attention that consumers devote to specific attributes during repeated conjoint choice tasks has been shown to improve preference measurements (Meißner et al. 2016; Yang et al. 2015).

Yet, what is still largely unknown is how eye movements, attention, and the utility of brands during choice are linked. Specifically, two key questions are (1) how trajectories of attention to each of the brands during the choice task contribute to the accumulation of utility and final choice, and (2) which fundamental attention processes contribute to the accumulation of utility and brand choice. Answering these questions is one step towards understanding the fundamental and possibly neurological links between attention, utility, and choice (Manohar and Husain 2013), and towards the more realistic, descriptive consumer choice theories that have been called for (Busemeyer and Johnson 2004; Caplin and Dean 2015; Krajbich et al. 2010; Stüttgen, Boatwright, and Monroe 2012; Willemsen et al. 2011).

This chapter follows up on this. It tests the hypothesis derived from RIT that attention during a complex choice task, the utility for each of the choice options, and final choice are closely aligned. We use eye-tracking and choice data from a study with a representative sample of 342 regular consumers who made a complex choice for one of five brands of smartphones on a realistic comparison website. We estimate a new, generalized Sequential Sampling Model (referred to as gSSM) (Forstmann, Ratcliff, and Wagenmakers 2016; Otter et al. 2008) to describe the relationship between consumers' eye movements during the choice task, trajectories of attention to each of the brands, and final brand choice. We find that over and above the total sum of attention for each brand, the temporal trajectories of attention during the choice task contribute to brand utility. Moreover, we find that over and above trajectories of the quantity of attention (reflected in eye fixations), trajectories of integration attention (reflected in within-brand saccades), contribute to brand utility and choice.

In addition, we use model estimates to infer the accumulation of brand utility during the task and find that attention already marks the final brand choice well before it is implemented. In fact, the model correctly inferred brand choice of 56% of consumers halfway before choice was expressed. Model performance rose to 77% one quarter before choice implementation, with a final hit rate of 85%. This reveals a much earlier attention bias effect for the chosen brand than what has been reported before. We also find that attention trajectories, rather than mere inertia or habits, account for state dependence effects in brand choice. Finally, the information-density of the decision-environment influences the level of attention that consumers devote to making a decision, but not the trajectories of attention, nor the association between attention and brand choice. Taken together, these findings support the fundamental link between eye movements, attention and utility accumulation during brand choice.

The next section presents our model and hypotheses. Then we describe our data, the econometric specification of the model, and the estimation results. The final section offers implications of the findings for consumer choice theory and for practice.

2.2 Theory

Consumers move their eyes when making a choice between multiple alternatives on a visual display, such as a comparison website. These eye movements comprise fixations and saccades. During a fixation, the eye is relative still (for about 200-400 msec.) and the gaze is directed to a specific location in space to acquire information from it. Because visual acuity rapidly drops-off with increasing distance from the center of the gaze, people need to move their eyes in order to acquire information from different locations in the visual display (van der Lans et al. 2008b). During such saccades, the gaze is rapidly redirected (20-50 msec.), while vision is actively suppressed to prevent blurring (Hutton 2008).

We propose a descriptive model of the relationship between eye movements that consumers make during a choice task and the accumulation of brand utilities. It is a generalized Sequential Sampling Model (gSSM) (Forstmann et al. 2016; Otter et al. 2008). It specifies that observable, overt eye movements that consumers use to sample information from a visual display with choice options reflect unobservable, covert attention processes that contribute to the accumulation of utility for the choice options. Thus, trajectories of covert attention (θ) connect overt eye movements (y) to the accumulation of brand utilities (u). In this way, the model and research are part of a broader effort to describe choice behavior and preference formation when the determinant processes are intrinsically unobservable (Caplin and Dean 2015). We first present the basic model and in a next section describe the econometric, restricted, model that we estimate. Specifically:

$$y_{ibt}^{g} = \theta_{ibt}^{g} + \xi_{ibt}^{g}, \tag{1}$$

$$u_{ibt} = \beta_t^{g} \theta_{ibt}^{g} + \varepsilon_{ibt}. \tag{2}$$

There are two crucial links, namely between overt eye movements and covert attention in eq. 1, and between covert attention and utility in eq. 2. Eq. 1 specifies that g types of observable eye movements (y_{ibt}^{g}, described later) of consumer i to brand b at time t during a choice task are a function of the covert attention of key interest (θ_{ibt}^{g}) and various sources of unobserved heterogeneity (ξ_{ibt}^{g}). Eq. 2 specifies two sources of utility (u_{ibt}) of brand b for consumer i at time t during a choice task, namely the weighted covert attention to the brand ($\beta_t^{g} \theta_{ibt}^{g}$), and unobserved, exogenous shocks to utility (ε_{ibt}). The ξ_{ibt}^{g} and ε_{ibt}

components reflect the notion that the information signals that consumers obtain about the utility of brands during search come with error (Maćkowiak et al. 2018). Model features are next.

2.2.1 Eye Movements and Attention

Eye movements of consumers during decision-making tasks are overt measures of the covert attention processes that take place during these tasks (Glaholt and Reingold 2011; Orquin and Loose 2013). Yet, eye movements are fallible indicators of the attentional processes of key interest for three reasons. First, covert attention and overt eye movements can be dissociated at a specific point in time, for instance to maintain a smooth flow of information processing. Then, like a rubber band, consumers' attention can already move to a location in space where it expects certain information before the eyes move, and the eyes can already move due to a salient event in the visual periphery before attention does (Hutton 2008). Second, the neural systems in the visual brain that direct the eye to locations in space proceed with some fixation-location error. In that case, eye saccades miss the intended exact x-y location in space, which may require corrective eye movements (Hutton 2008; Reichle and Drieghe 2015). Third, the recording of eye movements by eye-trackers proceeds with error. For common commercial eye-trackers such measurement error is small at .5 degrees of visual angle or less. Yet, that error is non-ignorable and varies between people and stimuli (van der Lans, Wedel, and Pieters 2011).

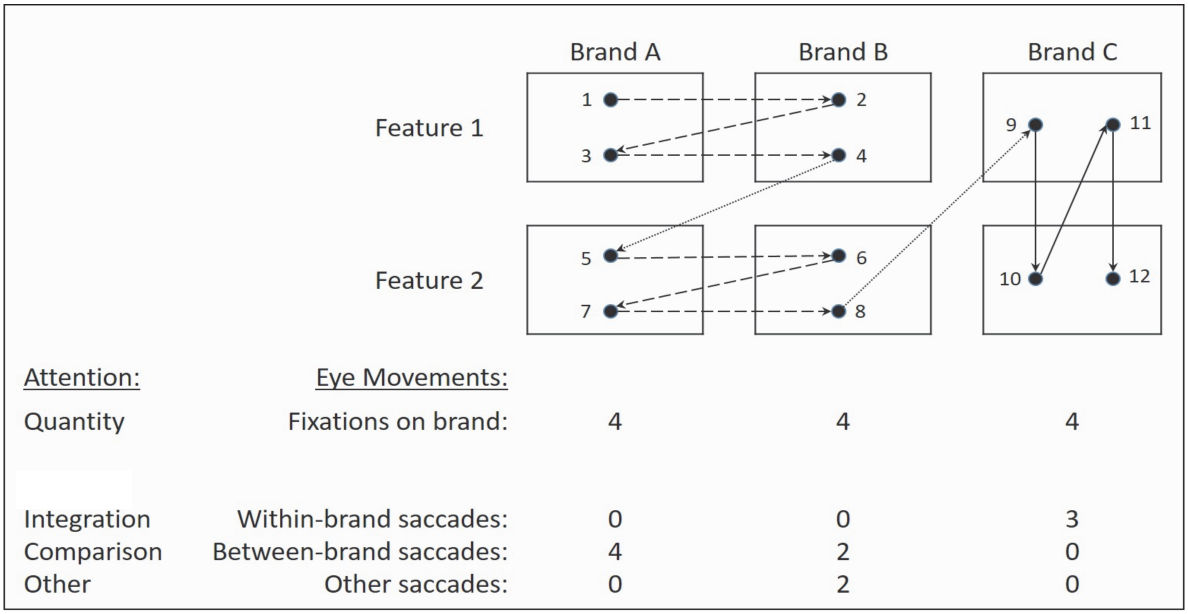
The model and experiment account for measurement error in eye movements in two ways. First, we aggregate overt eye-movement measures into larger areas-of-interest than their exact x-y location in space and into larger time bins than single fixations, in line with prior research (Pieters and Warlop 1999). For instance, Meißner et al. (2016) aggregated eye movements across a display of coffee machines into six features of three brands and into two time bins, namely the first and last 50%. Second, we decompose overt eye movements that consumers make for each of the brands in the choice set into covert attention (θ, in eq. 1) and measurement error (ξ). Shi et al. (2013) decomposed patterns of overt saccades that consumers made for the choice set as a whole into covert attentional strategies and measurement error. Our model builds on this by examining fixations and saccades for each brand in the choice set.

2.2.2 Quantity and Type of Attention

An eye fixation indicates whether or not attention has been devoted to a particular area-of-interest in a visual display, such as a brand. The number of eye fixations reflects the quantity of attention for the brand, which has been shown to predict overall liking of the brand, consideration, and choice (Chandon et al. 2009).

Eye saccades indicate specific types of attention to brands (Arieli et al. 2011; Lohse and Johnson 1996; Pieters and Warlop 1999; Shi et al. 2013). We propose that these specific types of attention contains additional information about the utility of brands, over and above the information contained in the mere quantity of attention. Figure 2.1 presents this idea. It displays a hypothetical pattern of eye movements across a visual display with three brands (A to C) each with two features (1 and 2). Three fundamental types of eye saccades are (a) between different features of the same brand, labelled “within-brand saccades”, (b) between the same feature of two different brands, labeled “between-brand saccades”, and (c) between different features of different brands, labelled “other saccades.” In Figure 2.1, the saccade between fixations 11 and 12 is within brand C, the saccade between fixations 1 and 2 is between brands A and B, and the saccade between fixations 8 and 9 is from the “other” category.

Figure 2.1 Attention and eye movements during brand choice



Within-brand saccades reflect attention to integrate information about a single brand into an overall judgment or evaluation, while between-brand saccades reflect attention to compare information across brands in order to learn about their performance. Such attention for, respectively, integration or comparison has been likened to foraging for value or foraging for information (Manohar and Husain 2013), value construction or value encoding (Willemsen et al. 2011), and holistic or component information processing (Arieli et al. 2011).

Importantly, the same quantity of attention, reflected in eye fixations, can be due to qualitatively different attention processes, as reflected in eye saccades. Figure 2.1 illustrates this. Saccades are assigned to the brand they originate from, where the decision to move the eye is made (Hutton 2008). Each of the three brands in Figure 2.1 receives four eye fixations. If only the quantity of attention would contribute to utility, the three brands would have the same choice likelihood. Yet, whereas brand A and B receive, respectively, 4 and 2 between-brand saccades which reflect attention for comparison, brand C receives 3 within-brand saccades which reflect attention for integration. If the attention type contributes to utility, choice probabilities of brands depend on the importance weight of the attention types. Specifically, attention for comparison aims to assess the performance of brands vis-à-vis each other (Arieli et al. 2011; Willemsen et al. 2011). In contrast, attention for information integration aims to assess whether the performance of brands on various features outweighs their costs and relative weaknesses (Manohar and Husain 2013). The "other" attention type most likely fulfills attentional "bookkeeping" functions such as searching the display for new information, and shifting between different attentional strategies (Shi et al. 2013) that are less central to the accumulation of utility. The model in eqs. 1 and 2 allows both attention quantity and type to contribute to brand utility, as indicated by superscript g.

2.2.3 Utility Accumulation and Choice

The model in eqs. 1 and 2 lets covert attention θ_{ibt}^{g} be brand-specific, as indicated by subscript b, and time-varying during the choice task, as indicated by the subscript t. This is supported by evidence that the eventually chosen option receives progressively more attention towards the end of a choice task (Fiedler and Glöckner 2012; Glaholt and Reingold 2011; Meißner et al. 2016; Willemsen et al. 2011). For instance, Atalay et al. (2012) found that the frequency of fixations on the finally chosen brand of vitamins and food-replacement bars rose in the final five seconds before choice. This is consistent with evidence that consumers transit through different stages when making a decision and that attention during these stages has different functions (Gidlöf et al. 2013; Shi et al. 2013; Stüttgen et al. 2012; Willemsen et al. 2011). Russo and Leclerc (1994) identified three stages in a study on choice for supermarket products such as peanut butter. In an early, orientation stage, consumers use eye movements to inspect products and learn about their content. In the following evaluation stage, consumers use eye movements to form an overall judgement or evaluation of specific brands. In a final, brief verification stage, eye movements aimed to ensure that important information for brand evaluation was not missed. This suggests that later attention to evaluate carries more weight than earlier attention to inspect (Willemsen et al. 2011).

2.2.4 Predictions and Contribution

The gSSM specifies the relationship between eye movements of consumers on brands during a choice task and the accumulation of utility of the brands. It posits covert trajectories of the quantity and type of attention as the link between overt eye movements and the accumulation of utility of the brands. Consistent with other SSMs (Krajbich et al. 2010) and RIT (Caplin and Dean 2015; Matějka and McKay 2015), it specifies that attention to brands in a visual display is biased towards the brand that is eventually chosen, rather than being uniformly distributed across brands. It extends prior work in three important ways, and makes novel predictions.

First, the model identifies covert attention and measurement error from overt eye movements. This improves on earlier work which rests on the assumption that eye movements are error-free attention measures (Atalay et al. 2012; Chandon et al. 2009; Krajbich et al. 2010; Meißner et al. 2016; Pieters and Warlop 1999; Reutskaja et al. 2011). As covert attention should be a less noisy indicator of utility than overt eye movements are, we predict the former to contribute more to brand utility (Hypothesis 1).

Second, the model identifies attention quantity, reflected in eye fixations, and attention types, reflected in eye saccades, and allows both to contribute to brand utility. This extends earlier work which rests on the assumption that eye fixations contain the key or all information about brand utility (Atalay et al. 2012; Chandon et al. 2009; Krajbich et al. 2010). Identifying the contribution of various types of attention to brand utility has been identified as an important area for new choice theories and research (Krajbich et al. 2010; Meißner et al. 2016, p. 16), but is still largely unexplored (an exception is Shi et al. (2013)). We predict that over and above attention quantity, the attention type contributes to brand utility (Hypothesis 2a). And more specifically, we predict that attention for information integration contributes more to brand utility than attention for comparison, and other attention do (Hypothesis 2b).

Third, the model specifies that attention trajectories and their contribution to brand utility systematically vary across the time course of the choice task. The term $\beta_t^g \theta_{ibt}^g$ in eq. 2 is a generalized form of the drift rate in SSM (Krajbich et al. 2010), which is the mean rate of change over time of the value of choosing an option or not. Our formulation extends prior work which rests on the assumption that the attention share of brands in the choice set is constant over time (Chandon et al. 2009; Pieters and Warlop 1999) or that it changes over time but that its contribution to brand utility is constant (Atalay et al. 2012; Krajbich et al. 2012; Meißner et al. 2016). If attention and its contribution to utility were time-invariant, differences between brands in accumulated (summed) attention at the end of the choice task would fully capture their utility differences. Instead, our theory predicts that the *trajectories* of attention to the brands during the choice task capture the accumulation of brand utility. More specifically, we predict that later attention contributes more to brand utility than earlier attention does (Hypothesis 3a). This is in line with neural evidence (Shimojo et al. 2003) that the accumulation of information about the utilities of choice options is reflected in an increased attention bias towards the finally chosen brand. H2a, H2b and H3a imply that both the later quantity and type of attention contribute more to brand utility than earlier attention does, and that in particular attention for information integration does so (Hypothesis 3b).

Finally, RIT also suggests that stickiness in choices is due to inattention to alternative choice options rather than being due to choice inertia or habitual choice (Steiner et al. 2017). Then, current brand ownership should affect choice and attention trajectories but, conditional on those, should have no direct effect on choice. Specifically, we test if stickiness in choices indeed arises due to biased attention to the owned brand (Hypothesis 4).

2.3 Data

2.3.1 Background and Sample

Our study simulates a typical on-line product comparison situation in which consumers evaluate a set of smartphones, from now on called devices, and choose one. Participants were presented with a side-by-side comparison of five devices, on 24-inch TFT computer monitors, as is common on many wireless carrier, retailer, and reviewer websites. The choice set consisted of the Apple iPhone 5 (brand A), Samsung Galaxy Note II (brand B), Nokia Lumia 920 (brand C), HTC One (brand D), Motorola Droid Razr Maxx HD (brand E). These were the most common devices in on-line product reviews and the most recent versions of each brand at the time of data collection (Spring 2013). Study participants were instructed to review the presented information about the devices and chose the device that they would be most likely to purchase.

Tobii Insight, a dedicated eye-tracking and market research firm, conducted sampling and data collection for the study (https://www.tobiipro.com/insight/). It drew a stratified sample of consumers who had indicated to be in the market for a new smartphone, from large, locally representative participant pools, from three locations in the continental US (Washington DC, Cincinnati, and San Diego). Data collection took place in dedicated research areas in shopping centers in each of the three locations.

The sample comprises 342 participants. Stratification ensured representation of four user groups: users of the two leading brands (30% brand A, and 24% brand B), other brand users (25%), and current non-device users (21%) were represented in the final sample. In the sample, 47% was female, and 30% was between 18-29 years, 39% between 30 and 49 years, 20% between 50 and 65, and 11% was over 65 years. Seventy-two percent of the sample (246 of 342) currently owned a smartphone of one of the five brands in the display, respectively 101 Apple, 83 Samsung, 38 HTC, 22 Motorola, and 2 Nokia. On average, participants indicated a 71% likelihood (0 to 100% scale) of purchasing a new smartphone in the next nine months. Participants received $50 to cover transportation costs of commuting to one of the data collection facilities and volunteering their time.

2.3.2 Design and Stimuli

To account for the possibility that the links between eye movements, attention, and brand utility depend on the information density in the choice task, we experimentally manipulated that in three levels (low, medium, and high). Information density or "complexity" was manipulated, between-participants, by varying the number of features presented to the participants on the computer screens (18 in low, 29 in medium, and 39 in high). Note that the low information density condition here is still denser than common in choice research (e.g., 3 brands and 6 features: Meißner et al. (2016); 4 brands and 12 attributes: Shi et al. (2013); 7 brands and 7 attributes: Lohse and Johnson (1996)). Participants were randomly assigned to one of the conditions (respective *n*s are 113 in low, 118 in medium, and 111 in high). Devices were shown in the columns and features in the rows, with the device name/model, colors, and the price category always displayed at the top of the page, as is common. Figure 2.2 provides examples from the low and high information conditions.

2.3.3 Eye Movements and Brand Choice

Eye-movement recording was done with Tobii 60XL infra-red eye-trackers integrated in 24 inch TFT computer monitors (screen resolution: 1920 x 1200 pixels), using sampling rate of 60hz, with a typical accuracy of 0.5 degrees of visual angle. Participants were free to move their head in a virtual box of 44 cm width x 22 cm height.

Figure 2.2 Sample choice displays

Low Information

<table>
<tr><th></th><th></th><th>Apple iPhone 5</th><th>HTC One</th><th>Samsung Galaxy Note II</th><th>Nokia Lumia 920</th><th>Motorola Droid RAZR MAXX HD</th></tr>
<tr><td></td><td>Colors Available</td><td>Click to Buy</td><td>Click to Buy</td><td>Click to Buy</td><td>Click to Buy</td><td>Click to Buy</td></tr>
<tr><td>Price</td><td>Price with 2-year Contract</td><td>$249.99</td><td>$199.99</td><td>$249.99</td><td>$99.99</td><td>$199.99</td></tr>
<tr><td>Wireless Capabilities</td><td>Band and Mode</td><td>|EDGE|GSM|Quad-band GSM|3G|UMTS|Quad-band UMTS|HSPA+|HSDPA|CDMA|4G LTE|Dual-band LTE|</td><td>|EDGE|GSM|Quad-band GSM|GPRS|3G|UMTS|HSDPA |CDMA|4G|4G LTE|Dual-band LTE|</td><td>|GSM|3G|UMTS|HSPA+|HSDPA |CDMA|4G LTE|Dual-band LTE|Quad-band|</td><td>|EDGE|GSM|Quad-band GSM|3G|Quad-band UMTS|HSPA+|HSDPA|4G|4G LTE|Dual-band LTE|</td><td>|EDGE|GSM|Quad-band GSM|3G|UMTS|HSPA+|HSDPA |CDMA|4G LTE|</td></tr>
<tr><td></td><td>Bluetooth</td><td>4.0</td><td>4.0</td><td>4.0</td><td>3.0 + HS</td><td>4.0</td></tr>
<tr><td></td><td>WiFi</td><td>802.11 a/b/g/n</td><td>802.11 a/b/g/n/ac</td><td>802.11 a/b/g/n</td><td>802.11 a/b/g/n</td><td>802.11 a/b/g/n</td></tr>
<tr><td>Operating System (OS) and Other Features</td><td>Mobile Operating System</td><td>iOS 6</td><td>Android 4.1 Jelly Bean</td><td>Android 4.1 Jelly Bean</td><td>Windows Phone 8</td><td>Android 4.1 Jelly Bean</td></tr>
<tr><td></td><td>Supported Email</td><td>POP3, IMAP, Push email (Exchange, Gmail)</td><td>POP3, IMAP, Push email (Exchange, Gmail, Hotmail)</td><td>POP3, IMAP, Push email (Exchange, Gmail)</td><td>SMTP, POP3, IMAP4, Push email (Exchange, Gmail, Hotmail)</td><td>POP3, IMAP, Push email (Exchange, Gmail)</td></tr>
<tr><td>Size</td><td>Product Dimensions</td><td>4.87" x 2.31" x 0.30"</td><td>5.41" x 2.69" x 0.37"</td><td>5.95" x 3.17" x 0.37"</td><td>5.13" x 2.79" x 0.42"</td><td>5.19" x 2.67" x 0.37"</td></tr>
<tr><td></td><td>Product Weight</td><td>3.95 oz.</td><td>5.04 oz.</td><td>6.44 oz.</td><td>6.53 oz.</td><td>5.54 oz.</td></tr>
<tr><td>Display</td><td>Screen Size</td><td>4.0 inches</td><td>4.7 inches</td><td>5.5 inches</td><td>4.5 inches</td><td>4.7 inches</td></tr>
<tr><td>Battery</td><td>Standby Time</td><td>Up to 9 days</td><td>Up to 20 days</td><td>Up to 19.5 days</td><td>Up to 13 days</td><td>Up to 16 days</td></tr>
<tr><td></td><td>Talk Time</td><td>Up to 8 hours</td><td>Up to 18 hours</td><td>Up to 21 hours</td><td>Up to 7.5 hours</td><td>Up to 32 hours</td></tr>
<tr><td>Camera</td><td>Camera Resolution</td><td>8.0MP</td><td>HTC UltraPixel 4MP Camera</td><td>8.0MP</td><td>8.7MP</td><td>8.0MP</td></tr>
<tr><td>Memory</td><td>Internal Memory</td><td>32GB</td><td>32GB</td><td>16GB</td><td>32GB</td><td>32GB</td></tr>
<tr><td></td><td>Memory Card</td><td>No</td><td>No</td><td>microSD up to 64GB</td><td>No</td><td>microSD up to 32GB</td></tr>
<tr><td></td><td>RAM</td><td>1GB</td><td>2GB</td><td>2GB</td><td>1GB</td><td>1GB</td></tr>
</table>

High Information

<table>
<tr><th></th><th></th><th>Motorola Droid RAZR MAXX HD</th><th>Apple iPhone 5</th><th>Samsung Galaxy Note II</th><th>HTC One</th><th>Nokia Lumia 920</th></tr>
<tr><td></td><td>Colors Available</td><td>Click to Buy</td><td>Click to Buy</td><td>Click to Buy</td><td>Click to Buy</td><td>Click to Buy</td></tr>
<tr><td>Price</td><td>Price with 2-year Contract</td><td>$199.99</td><td>$249.99</td><td>$249.99</td><td>$199.99</td><td>$99.99</td></tr>
<tr><td>Battery</td><td>Battery Type</td><td>Lithium-ion Polymer</td><td>Lithium-ion</td><td>Lithium-ion</td><td>Lithium-polymer</td><td>Lithium-polymer</td></tr>
<tr><td></td><td>Standby Time</td><td>Up to 16 days</td><td>Up to 9 days</td><td>Up to 19.5 days</td><td>Up to 20 days</td><td>Up to 13 days</td></tr>
<tr><td></td><td>Talk Time</td><td>Up to 32 hours</td><td>Up to 8 hours</td><td>Up to 21 hours</td><td>Up to 18 hours</td><td>Up to 7.5 hours</td></tr>
<tr><td>Display</td><td>Pixel Density</td><td>313 PPI</td><td>326 PPI</td><td>267 PPI</td><td>468 PPI</td><td>332 PPI</td></tr>
<tr><td></td><td>Screen Resolution</td><td>1280 x 720</td><td>1136 x 640</td><td>1280 x 720</td><td>1920 x 1080</td><td>1280 x 768</td></tr>
<tr><td></td><td>Screen Size</td><td>4.7 inches</td><td>4.0 inches</td><td>5.5 inches</td><td>4.7 inches</td><td>4.5 inches</td></tr>
<tr><td></td><td>Screen Type</td><td>OLED (Active, Color, Backlit)</td><td>LCD (Active, Color, Backlit)</td><td>OLED (Active, Color, Backlit)</td><td>LCD (Active, Color, Backlit)</td><td>LCD (Active, Color, Backlit)</td></tr>
<tr><td></td><td>Touch Screen</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td>Wireless Capabilities</td><td>Band and Mode</td><td>|EDGE|GSM|Quad-band GSM|3G|UMTS|HSPA+|HSDPA |CDMA|4G LTE|</td><td>|EDGE|GSM|Quad-band GSM|3G|UMTS|Quad-band UMTS|HSPA+|HSDPA|CDMA|4G LTE|Dual-band LTE|</td><td>|GSM|3G|UMTS|HSPA+|HSDPA |CDMA|4G LTE|Dual-band LTE|Quad-band|</td><td>|EDGE|GSM|Quad-band GSM|GPRS|3G|UMTS|HSDPA |CDMA|4G|4G LTE|Dual-band LTE|</td><td>|EDGE|GSM|Quad-band GSM|3G|Quad-band UMTS|HSPA+|HSDPA|4G|4G LTE|Dual-band LTE|</td></tr>
<tr><td></td><td>Bluetooth</td><td>4.0</td><td>4.0</td><td>4.0</td><td>4.0</td><td>3.0 + HS</td></tr>
<tr><td></td><td>Built-In GPS</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td></td><td>WiFi</td><td>802.11 a/b/g/n</td><td>802.11 a/b/g/n</td><td>802.11 a/b/g/n</td><td>802.11 a/b/g/n/ac</td><td>802.11 a/b/g/n</td></tr>
<tr><td>Memory</td><td>Internal Memory</td><td>32GB</td><td>32GB</td><td>16GB</td><td>32GB</td><td>32GB</td></tr>
<tr><td></td><td>Memory Card</td><td>microSD up to 32GB</td><td>No</td><td>microSD up to 64GB</td><td>No</td><td>No</td></tr>
<tr><td></td><td>RAM</td><td>1GB</td><td>1GB</td><td>2GB</td><td>2GB</td><td>1GB</td></tr>
<tr><td>Camera</td><td>Camera Resolution</td><td>8.0MP</td><td>8.0MP</td><td>8.0MP</td><td>HTC UltraPixel 4MP Camera</td><td>8.7MP</td></tr>
<tr><td></td><td>Records Video</td><td>1080p Full HD</td><td>1080p Full HD</td><td>1080p Full HD</td><td>1080p Full HD</td><td>1080p Full HD</td></tr>
<tr><td></td><td>Secondary Camera</td><td>1.3MP</td><td>1.2MP</td><td>1.9MP</td><td>2.1MP</td><td>1.2MP</td></tr>
<tr><td>Size</td><td>Product Dimensions</td><td>5.19" x 2.67" x 0.37"</td><td>4.87" x 2.31" x 0.30"</td><td>5.95" x 3.17" x 0.37"</td><td>5.41" x 2.69" x 0.37"</td><td>5.13" x 2.79" x 0.42"</td></tr>
<tr><td></td><td>Product Weight</td><td>5.54 oz.</td><td>3.95 oz.</td><td>6.44 oz.</td><td>5.04 oz.</td><td>6.53 oz.</td></tr>
<tr><td>Operating System (OS) and Other Features</td><td>CPU</td><td>1.5 GHz Dual Core</td><td>Not Specified by manufacturer</td><td>1.6 GHz Quad Core</td><td>1.7 GHz Quad Core</td><td>1.5 GHz Dual Core</td></tr>
<tr><td></td><td>Media - Audio</td><td>AAC, AAC+, AMR-NB, AMR-WB, eAAC+, MIDI, MP3, OGG, WMA v9, WMA v10</td><td>AAC, HE-AAC, MP3, MP3 VBR, Audible, Apple Lossless, AIFF, and WAV</td><td>MP3, OGG, WMA, AAC, ACC+, eAAC+, AMR-NB, AMR-WB, MIDI, WAV, AC-3, Flac</td><td>AAC, AMR, OGG, M4A, MID, MP3, WAV, WMA</td><td>MP3, QCELP, AMR-WB, AMR-NB, WMA 10 Pro, WMA 9, G.711, AAC-LC, AAC+/HEAAC, eAAC+/HEAACv2, ASF, MP4, AAC, M4A, 3GP, 3GZ</td></tr>
<tr><td></td><td>Media - Video</td><td>MPEG-4, H.263, H.264, VC-1, VP8</td><td>h.264 / AVC, Motion JPEG, MPEG-4, Quicktime</td><td>MPEG4, H.263, H.264, VC-1, DivX, WMV, VP8, 3GP(MP4), AVI, FLV, MKV, WebM</td><td>3GP, 3G2, MP4, WMV, AVI</td><td>H.264 / AVC, MPEG-4, VC-1, Windows video, H.263, WMV, AVI, 3GP, 3G2, M4V, MOV</td></tr>
<tr><td></td><td>MMS</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td></td><td>Mobile Operating System</td><td>Android 4.1 Jelly Bean</td><td>iOS 6</td><td>Android 4.1 Jelly Bean</td><td>Android 4.1 Jelly Bean</td><td>Windows Phone 8</td></tr>
<tr><td></td><td>NFC</td><td>Yes</td><td>No</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td></td><td>OS Support</td><td>Windows (7, Vista, XP), Mac OS</td><td>Windows (7, Vista, XP), Mac OS (OS X)</td><td>Windows (7, Vista, XP), Mac OS</td><td>Windows (7, Vista, XP), Mac OS</td><td>Windows (7, Vista, XP), Mac OS</td></tr>
<tr><td></td><td>Phone Style</td><td>Bar phone</td><td>Bar phone</td><td>Bar phone</td><td>Bar phone</td><td>Bar phone</td></tr>
<tr><td></td><td>QWERTY Keyboard</td><td>No</td><td>No</td><td>No</td><td>No</td><td>No</td></tr>
<tr><td></td><td>Released (US)</td><td>10/19/2012</td><td>9/21/2012</td><td>11/9/2012</td><td>4/19/2013</td><td>11/9/2012</td></tr>
<tr><td></td><td>Speakers</td><td>Mono</td><td>Mono</td><td>Mono</td><td>Stereo</td><td>Mono</td></tr>
<tr><td></td><td>Supported Email</td><td>POP3, IMAP, Push email (Exchange, Gmail)</td><td>POP3, IMAP, Push email (Exchange, Gmail)</td><td>POP3, IMAP, Push email (Exchange, Gmail)</td><td>POP3, IMAP, Push email (Exchange, Gmail, Hotmail)</td><td>SMTP, POP3, IMAP4, Push email (Exchange, Gmail, Hotmail)</td></tr>
<tr><td></td><td>TV-out</td><td>Micro-HDMI, Wireless with DLNA</td><td>Lightning Digital AV Adapter, Lightning to VGA Adapter, Wirelessly</td><td>MHL / HDMI (micro-USB)</td><td>MHL / HDMI (micro-USB), Wireless with DLNA</td><td>Wireless with DLNA</td></tr>
<tr><td></td><td>Voice Dial</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
<tr><td></td><td>Warranty (Labor/Parts)</td><td>1 Year</td><td>1 Year</td><td>1 Year</td><td>1 Year</td><td>1 Year</td></tr>
</table>

Eye fixations were determined from the raw eye-movement recordings using default factory settings. Eye movements were aggregated for each participant in two ways. Fixation locations were aggregated into areas-of-interest covering each a feature for a brand in the visual display. The sequence of eye fixations until choice was aggregated into four time bins for each participant, labelled Quarters 1 to 4, respectively, 0-25%, 26-50%, 51-75%, and 76-100% of eye fixations. These periods cover common stages of information exploration (Q1), evaluation (Q2 and Q3), and verification and choice (Q4) (Gidlöf et al. 2013; Russo and

Leclerc 1994; Willemsen et al. 2011).

With 342 participants, 5 brands, 4 time periods, and 4 eye-movement measures, the dataset contains 27,360 observations. On average, participants made 236 eye fixations (SD = 200) before making their choice. Information condition influenced fixation frequencies until choice was made ($F(2, 333) = 7.56$, $p < .001$): participants in the low information condition fixated less (M = 185, SD = 122) than participants in the medium (M = 239, SD = 185) and high condition (M = 285, SD = 259), who did not differ from each other. Of the 342 participants in the sample, 129 chose the brand they currently owned (“loyals”), 117 switched brands (“switchers”), and 96 did not own a device in the category yet (n = 71) or owned another brand than on display (n = 25) (“others”). Also, these customer segments differed in fixation frequency ($F(2, 333) = 7.47$, $p < .001$): loyals fixated less (M = 188, SD = 168) than switchers (M = 259, SD = 181) and other customers did (M = 274, SD = 246), who did not differ from each other. The interaction between information condition and customer segment was not significant ($F(4, 333) = 1.36$, $p = .24$). Choice shares were 26%, 29%, 8%, 21%, and 16% for, respectively, brands A to E. There were no differences in state dependence between brands ($\chi^2(4) = 2.16$, $p = .71$), and information condition did not affect the brand being chosen ($\chi^2(8) = 6.33$, $p = .61$).

2.4 Model Specification

The theoretical model in eqs. 1 and 2 specifies a link between eye movements, trajectories of covert attention, and the accumulation of utility of the choice options during a choice task. It (1) infers latent attention trajectories from observed eye-movement measures, and (2) quantifies the link between the attention trajectories and the accumulation of brand utilities as revealed in choice.

2.4.1 Attention Trajectories

We use a reduced form of eq. 1 to identify attention trajectories. Specifically, we decompose the eye-movement measures of consumers to each of the brands during the choice task into (1) three components of the attention trajectories and (2) unobserved heterogeneities. We further decompose each attention trajectory component into (1) the consumer-brand-specific part of prime interest, and (2) the consumer-specific part which is constant across brands. We use a general latent trajectory specification for this (Muthén 1997; Muthén et al. 2011):

$$y_{ibt}^{g} = \eta_{ib0}^{g} + \eta_{ib1}^{g} x_t + \eta_{ib2}^{g} x_t^2 + \xi_{i0t}^{g} + \xi_{ibt}^{g}, \tag{3}$$

$$\eta_{ibk}^{g} = \eta_{00k}^{g} + \theta_{i0k}^{g} + \theta_{ibk}^{g}. \tag{4}$$

In eqs. 3 and 4 superscript g indexes eye-movement measures, respectively brand fixations, within-brand saccades, between-brand saccades, and other saccades. Subscript t indexes time periods, b indexes brands, and i indexes consumers. The number of fixations and saccades are natural-log transformed (after adding 1 to accommodate zero frequencies: $y_{ibt}^{g} = ln(\breve{y}_{ibt}^{g} + 1)$) to normalize their distribution (Pieters and Wedel 2004; Rosbergen, Pieters, and Wedel 1997). Then, $g = 1, \dots, 4$, $t = 1, \dots, 4$, $i = 1, \dots, 342$, $b = 1, \dots, 5$.

In eq. 3, x_t and x_t^2 denote, respectively, linear change (coded as 0, 1, 2, 3 for the 4 time periods) and quadratic change (0, 1, 4, 9), as deviations from the initial level (η_{ib0}^{g}) of the attention trajectory, and k denotes the three attention trajectory components. Thus, eq. 3 specifies that the trajectory of each of the eye-movement measures is the weighted sum of three components of an attention trajectory, and time-varying unobserved heterogeneity for consumers (ξ_{i0t}^{g}), and brands and consumers (ξ_{ibt}^{g}). The three attention trajectory components are the initial level after the first period (η_{ib0}^{g}), a linear change component (η_{ib1}^{g}) and a quadratic change component (η_{ib2}^{g}). Then eq. 4 decomposes each of the attention trajectory components (η_{ibk}^{g}) into an overall mean (η_{00k}^{g}), consumer-specific attention (θ_{i0k}^{g}), and brand-consumer-specific attention, from now on called brand-specific (θ_{ibk}^{g}), for each of the four eye-movement measures.

2.4.2 Accumulation of Utility

We use a reduced form of eq. 2 to specify the contribution that the three components of the brand-specific attention trajectories have to the utility of the brands, while accounting for consumer prior states and market-level preferences:

$$u_{ib} = \alpha_b + \alpha_6 BO_{ib} + \beta_k^g \theta_{ibk}^g + \varepsilon_{ib}. \tag{5}$$

In eq. 5, u_{ib} is the overall utility of brand b for consumer i. The first two terms capture the effect of consumers' prior states and knowledge (Matějka and McKay 2015) independent of attention trajectories. The term α_b captures intrinsic market-level preferences for the brands in the choice set, using brand fixed-effects ($\alpha_1 = 0$, for identification). The term $\alpha_6 BO_{ib}$ captures state dependence effects (Dubé et al. 2008), where BO_{ib} indicates if consumer i currently owns brand b ($BO_{ib} = 1$ if consumer i owns brand j, and 0 otherwise), and α_6 reflects the size of the state dependence effect. These terms effectively control for market conditions and prior states that may influence brand utility independent of attention, and that might confound inferences of the contribution of attention to utility if left unaccounted for. The term $\beta_k^g \theta_{ibk}^g$ captures the contribution of the attention trajectories to brand utility. It allows different types of attention, as expressed in the superscript g, to contribute to brand utility and allows their contribution to be time-varying, as expressed in the subscript k.

The random components ε_{ib} are assumed to be type I extreme value distributed, which gives a conditional logit formulation of brand choice (McFadden 1973). Specifically:

$$p(BC_i = b|\alpha, \beta) = \frac{exp\,(u_{ib})}{\sum_{l=1}^{5} exp(u_{il})}, \tag{6}$$

with $p(BC_i = b)$ the probability that consumer i chooses brand b. From eqs. 5 and 6 we can use backward induction to infer the accumulated brand utilities in each period as:

$$\hat{u}_{ibt} = \hat{\alpha}_b + \hat{\alpha}_6 BO_{ib} + w_{kt}\hat{\beta}_k^g \hat{\theta}_{ibk}^g. \tag{7}$$

In eq. 7, the weight $w_{kt} = \left(\frac{t-1}{3}\right)^k$ transforms the contribution ($\hat{\beta}_k^g \hat{\theta}_{ibk}^g$) of the k (0 to 2) attention trajectories components to brand choice into brand utilities brand ($\hat{u}_{ibt}$) in each period t, while controlling for prior consumer states and knowledge. We use these estimated utilities to examine how brand choice probabilities change during a single brand choice.

2.4.3 Determinants of Attention Trajectories

The formulation in eq. 5 lets consumer prior states and market-level preferences have direct effects on utility. Our model accommodates these and additional effects also on the attention trajectories in eq. 5. This expresses that attention trajectories are determined exogenously by stimulus and endogenously by person characteristics (Bordalo, Gennaioli, and Shleifer 2013; Chandon et al. 2009; Pieters and Warlop 1999). Specifically:

$$\theta_{i0k}^g = \gamma_{k1}^g IC_i + \gamma_{k2}^g PO_i + r_{i0k}^g, \tag{8}$$

$$\theta_{ibk}^g = \gamma_{k3b}^g BP_b + \gamma_{k4}^g BO_{ib} + r_{ibk}^g. \tag{9}$$

Eq. 8 quantifies the extent to which consumer-specific attention trajectories are influenced by the information condition in the experiment (IC_i; coded -1 = low, 0 = medium, and 1 = high) and current product ownership ($PO_i = 1$ if consumer i owns a brand in the product category, and 0 otherwise), and consumer-specific heterogeneity (r_{i0k}^g). Eq. 9 quantifies how much brand-specific attention trajectories are influenced by market-level brand preferences (BP_b), current brand ownership (BO_{ib}), and brand-specific heterogeneity (r_{ibk}^g). Allowing market-level brand preferences and current brand ownership to influence both attention and choice reduces the possibility that estimated links between attention and choice are spurious. This is interesting of its own by documenting possible state dependence and status quo effects (Dubé et al. 2008) on attention and choice for complex products. Our research is the first to link attention in a brand choice context to previous brand choice, and the first to quantify the effects of information richness on attention trajectories.

Heterogeneities in attention trajectories (r_{i0}^{g} and r_{ib}^{g}) and eye-movement measures (ξ_{i0}^{g} and ξ_{ib}^{g}) are assumed normally distributed with mean zero and uncorrelated between the brand and consumer-levels. Components of the attention trajectories (k) for all eye-movement measures (g) are allowed to correlate at the brand ($r_{ib} \sim N(0, \Psi_1)$) and consumer-level ($r_{i0} \sim N(0, \Psi_0)$). Eye-movement measures are allowed to correlate at each time period: $\xi_{i0} \sim N(0, \Sigma_0), \xi_{ib} \sim N(0, \Sigma_1)$, with a block-diagonal structure on Σ_0 and Σ_1 at each time point (details in Appendix A).

2.4.4 Model Estimation

The joint likelihood of the model is:

$$\mathcal{L}(BC, Y|\Gamma) = \prod_{i=1}^{342} p(BC_i = b_i^*|\alpha, \beta) \prod_{b=1}^{5} p(y_{ib}|\eta_{ib}, \Sigma) p(\eta_{ib}|\eta_{00}, \gamma, \Psi) \,, \quad (10)$$

where BC is the vector of observed brand choices, Y contains the eye-movement measures, $\Gamma \equiv [\alpha, \beta, \eta_{ib}, \eta_{00}, \gamma, \Sigma, \Psi]$ denotes all the parameters of the model: choice weights (α, β), brand and consumer specific attention trajectories (η_{ib}), overall attention trajectories (η_{00}), brand and consumer specific effects (γ), and variances of the unobserved heterogeneities in eye-movement measures (Σ) and attention trajectories (Ψ) (details in Appendix A). The model is estimated in R using the RStan package (Carpenter et al. 2017; R Core Team 2018). We assess convergence using the Gelman-Rubin diagnostic – potential scale reduction factor (Gelman et al. 2013) and evaluate model fit based on the log marginal density (LMD) computed using the Gelfand and Dey method for the attention part (Gelfand and Dey 1994; Koop, Poirier, and Tobias 2007) and the Newton and Raftery method for the choice part (Rossi, Allenby, and McCulloch 2005), and the hit rate for brand choice (percentage correctly predicted choices). Models with LMDs closer to zero and with higher hit rates are preferable. One-tailed Bayesian *p*-values are reported for parameter estimates, indicating the percentage of posterior draws with a different sign from that of the estimated mean of the parameter.

2.5 Results

2.5.1 Attention Trajectories

Table 2.1 has descriptive information about eye movements during the choice task for the three conditions, and for the chosen and non-chosen brands. To facilitate interpretation, it presents the shares of the three types of brand eye-saccades (within-brand, between-brand, and other) rather than their raw frequencies. We confirmed that all three components of the attention trajectories are required to describe the eye-movement patterns, and that consumer-level and brand-level factors influence the attention trajectories by comparing the fit of the full model against three alternatives. A model with all three attention trajectory components and effects of current brand and product ownership, and information condition (LMD = -22,121) outperformed models with, respectively, only (1) the three attention trajectories (LMD = -24,005), (2) the initial level and linear component (LMD = -26,716), and (3) the initial level (LMD = -32,728). Table 2.2 provides estimates of the attention trajectories for the full model.

Consumer-level Effects. By construction (normalization in time bins), consumer-level trajectories of the quantity of attention are constant across the choice task, and can only differ in their overall level. As reported in the Data section (2.3.3), the three information density conditions influenced the total quantity of attention that consumers allocated during the choice task (γ_{11}^{1} = .19, *SD* = .06, p < .001, Table 2.2 top). Importantly, there were no differences between information conditions in integration, comparison, or other attention during the choice task (k = 1, k = 2). Product ownership did not significantly influence the consumer-level attention trajectories (all ps $\geq$ .14). Thus, attention trajectories of consumers varied in their level (k = 0) but not in their shape (k = 1, k = 2) across information conditions and prior product ownership. This lets us focus on the brand-level effects.

Table 2.1 Summary of eye movements

Eye Fixations and Shares of Saccades over Time	Information Density						Total		Choice-based			
	Low		Medium		High				Chosen Brand		Non-Chosen Brands	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1st Quarter:												
Eye fixation frequency	42.25	28.88	56.40	44.15	70.16	69.07	56.19	51.08	13.79	15.09	10.60	13.55
Within-brand saccade share	.38	.31	.35	.28	.37	.28	.37	.29	.44	.29	.35	.29
Between-brand saccade share	.37	.35	.40	.32	.38	.31	.39	.33	.33	.31	.40	.33
Other saccade share	.25	.29	.25	.25	.25	.25	.25	.26	.23	.24	.25	.27
2nd Quarter:												
Eye fixation frequency	47.91	32.93	60.80	49.7	70.71	66.32	59.76	52.04	14.88	13.66	11.22	13.17
Within-brand saccade share	.38	.30	.34	.27	.36	.28	.36	.28	.45	.28	.33	.28
Between-brand saccade share	.36	.32	.42	.31	.39	.31	.39	.31	.33	.28	.41	.32
Other saccade share	.25	.27	.24	.23	.25	.25	.25	.25	.22	.22	.25	.26
3rd Quarter:												
Eye fixation frequency	46.88	31.46	59.29	46.15	71.30	63.53	59.09	49.56	17.89	16.91	10.30	12.43
Within-brand saccade share	.38	.29	.37	.28	.37	.27	.38	.28	.50	.26	.34	.28
Between-brand saccade share	.38	.32	.38	.31	.36	.29	.37	.30	.31	.26	.39	.31
Other saccade share	.24	.25	.24	.25	.27	.26	.25	.25	.19	.16	.27	.27
4th Quarter:												
Eye fixation frequency	48.39	33.32	62.56	48.82	73.16	65.09	61.32	51.49	26.88	21.86	8.61	11.51
Within-brand saccade share	.41	.31	.41	.30	.38	.28	.40	.30	.60	.24	.34	.29
Between-brand saccade share	.33	.30	.34	.30	.34	.29	.33	.30	.21	.22	.37	.31
Other saccade share	.26	.27	.25	.26	.28	.27	.26	.27	.19	.18	.29	.29
Total:												
Eye fixation frequency	185.43	122.44	239.04	185.08	285.33	259.32	236.35	200.03	73.45	57.91	40.73	42.32
Within-brand saccade share	.39	.21	.36	.20	.37	.19	.37	.20	.50	.17	.34	.20
Between-brand saccade share	.37	.23	.40	.22	.37	.21	.38	.22	.29	.16	.40	.23
Other saccade share	.25	.18	.24	.17	.26	.16	.25	.17	.21	.12	.26	.18

Note: Total sample size 342 across 5 brands (342 x 5 = 1710), with 113 (565) in low, 118 (590) in medium, and 111 (555) in high information condition. For chosen brands n = 342, and for non-chosen brands n = 1368 (4 x 342). Average brand eye fixations are calculated across the four non-chosen brands.

Table 2.2 Attention trajectories

	Components of Attention Trajectories								
	Initial Level (k = 0)			Linear Change (k = 1)			Quadratic Change (k = 2)		
Predictors	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Consumer-level (i)									
Brand eye fixations (g = 1)									
Intercept (η^{g}_{00k})	2.09	.11	<.001	-.03	.07	.35	.00	.02	.49
Information density (γ^{g}_{k1})	.19	.06	<.001	-.04	.03	.09	.01	.01	.18
Product ownership (γ^{g}_{k2})	-.06	.11	.31	-.06	.06	.18	.00	.02	.48
Heterogeneity (r^{g}_{i0k})	.59	.06	<.001	.33	.07	<.001	.02	.01	<.001
Within-brand saccades (g = 2)									
Intercept (η^{g}_{00k})	1.27	.09	<.001	-.14	.09	.06	.04	.03	.06
Information density (γ^{g}_{k1})	.12	.05	.01	-.03	.04	.20	.01	.01	.33
Product ownership (γ^{g}_{k2})	-.10	.10	.14	.04	.08	.31	-.02	.02	.18
Heterogeneity (r^{g}_{i0k})	.30	.04	<.001	.35	.11	<.001	.03	.01	<.001
Between-brand saccades (g = 3)									
Intercept (η^{g}_{00k})	1.14	.09	<.001	.15	.09	.04	-.04	.03	.05
Information density (γ^{g}_{k1})	.18	.05	<.001	-.03	.04	.24	.00	.01	.42
Product ownership (γ^{g}_{k2})	.01	.10	.47	-.08	.08	.16	.00	.03	.44
Heterogeneity (r^{g}_{i0k})	.44	.05	<.001	.16	.04	<.001	.01	.00	<.001
Other saccades (g = 4)									
Intercept (η^{g}_{00k})	1.02	.08	<.001	-.03	.08	.37	.01	.02	.37
Information density (γ^{g}_{k1})	.14	.04	<.001	-.04	.04	.12	.01	.01	.12
Product ownership (γ^{g}_{k2})	-.06	.09	.25	-.04	.08	.28	.00	.02	.46
Heterogeneity (r^{g}_{i0k})	.31	.03	<.001	.11	.04	<.001	.01	.00	<.001
Brand-level (b)									
Brand eye fixations (g = 1)									
Brand ownership (γ^{g}_{k4})	.19	.06	<.001	.00	.08	.49	.04	.02	.04
Heterogeneity (r^{g}_{ibk})	.45	.06	<.001	.33	.07	<.001	.02	.01	<.001
Within-brand saccades (g = 2)									
Brand ownership (γ^{g}_{k4})	.20	.07	.003	.01	.09	.47	.04	.03	.12
Heterogeneity (r^{g}_{ibk})	.50	.08	<.001	.35	.11	<.001	.03	.01	<.001
Between-brand saccades (g = 3)									
Brand ownership (γ^{g}_{k4})	.06	.04	.07	.04	.07	.28	.00	.02	.42
Heterogeneity (r^{g}_{ibk})	.13	.03	<.001	.16	.04	<.001	.01	.00	<.001
Other saccades (g = 4)									
Brand ownership (γ^{g}_{k4})	.07	.05	.07	.04	.07	.26	.01	.02	.27
Heterogeneity (r^{g}_{ibk})	.17	.03	<.001	.11	.04	<.001	.01	.00	<.001

Note – *M* = Mean estimate; *SD* = Standard deviation; *p*-value = one-tailed Bayesian significance level. All eye-movement measures transformed natural log +1 prior to analysis. Information density: -1 = low, 0 = medium, 1 = high. Product -, and brand ownership: 1 = yes, 0 = no. Effects of market-level preferences (brand fixed-effects) not shown to save space. Heterogeneity refers to the variance of r^{g}_{i0k} and r^{g}_{ibk}.

Brand-level Effects. As hypothesized, current brand ownership influenced attention trajectories. The initial level of the quantity of attention to the currently owned brand was significantly higher than to the other brands (.19, SD = .06, $p < .001$, Table 2.2). This was mostly due to a higher level of integration attention to the currently owned brand (.20, SD = .07, $p = .003$) rather than to elevated levels of comparison (.06, SD = .04, $p = .07$) or other attention (.07, SD = .05, $p = .07$). Also, the currently owned brand attracted a higher attention quantity towards the end of the task (quadratic change .04, SD = .02, $p = .04$).

2.5.2 Contribution of Attention Trajectories to Brand Choice

The proposed gSSM specifies that the accumulation of brand utilities is a function of the time-varying weights (β_k^g) of time-varying, consumer- and brand-specific attention (θ_{ibk}^g).

Theory Tests and Model Selection. We compared our proposed model against five competing models. Table 2.3 summarizes the key assumptions of each model, indicates how their specification differs from the full model, and presents model fit results. We examine these two sources of utility, and start with the contribution of attention.

Model 1 has an LMD of -481 and a hit rate of 44%. This model assumes that attention during the choice task does not provide information about brand utility. Model 2 assumes that overt eye movements have a time-invariant contribution to brand utility, similar to previous research on attention and choice (Glaholt, Wu, and Reingold 2009; Krajbich et al. 2010; Pieters and Warlop 1999). Model 2 (LMD of -301 and hit rate of 68%) improves substantially over model 1, which provides initial support for our reasoning and RIT.

Models 3 and 4 specify that the trajectories of attention contain information about brand utilities. Model 3 assumes that only the trajectories of attention quantity, and model 4 that only those of the specific attention types contribute to brand utility. These models have similar fit (LMD of -170 and -175 respectively, and hit rate of 83%), both improving markedly over model 2. This reveals the importance of accounting for measurement error in

Table 2.3 Model selection

Model	Determinants of Brand Choice	Assumptions and Implications	Difference from Proposed Model	LMD (# pars)	Hit Rate
M1	Prior ownership and market-level preferences	Attention does not reflect brand utility.	$\beta_k^g = 0$, for $g = 1, \dots, 4$; and $k = 0, \dots, 2$	-481 (5)	44%
M2	M1 + sum of eye fixations	Sum of overt eye movements across all periods reflects brand utility.	$\theta_{ibk}^g = \sum_{t=1}^4 \breve{y}_{ibt}^g$, for $g = 1$ and $k = 1$ $\beta_k^g = 0$, for $g = 2, \dots, 4$ and $k = 0, \dots, 2$	-301 (6)	68%
M3	M1 + trajectories of attention quantity	Contribution of attention quantity to utility is time-varying. Utility is invariant to type of attention.	$\beta_k^g = 0$, for $g = 2, \dots, 4$; and $k = 0, \dots, 2$	-170 (8)	83%
M4	M1 + trajectories of attention type	Contribution of attention type to utility is time-varying. Utility is invariant to quantity of attention.	$\beta_k^g = 0$, for $g = 1$; and $k = 0, \dots, 2$	-175 (14)	83%
M5	M1 + initial and linear trajectory components of attention quantity and type	Increased attention during final stages of the choice task does not carry extra weight on brand utilities.	$\beta_k^g = 0$, for $g = 1, \dots, 4$; and $k = 2$	-267 (13)	69%
(M6) Proposed model		Trajectories of attention quantity and type contribute to brand utilities throughout the choice task.	Eqs. 3 to 5, 8, 9	-164 (17)	85%

Note: LMD is Log Marginal Density. Hit rate is the percent of correctly inferred brand choices. Null model without predictors has LMD -550 and hit rate 20% (1 out of 5). g = 1, …, 4 indexes eye-movement measures, respectively brand fixations, within-brand saccades, between-brand saccades, and other saccades; t = 1, …, 4 indexes time periods, b = 1, …, 5 indexes brands, and i = 1, …, 342 indexes consumers.

eye movements, and for the moment-to-moment changes in attention that take place during choice, which supports H1.

Model 5 assumes that trajectories of attention quantity and type contribute to brand utility, but that the contribution increases linearly over time, without a boost towards the end. It is a restricted version of the full model without the quadratic change component of the attention trajectories. It has an LMD of -267 and a hit rate of 69%, both much worse than the full model, and models 2 and 3 which contained the quadratic change component.

The full model outperforms all restricted models (LMD of -164 and hit rate of 85%). These results have two key implications. First, the findings support that the contribution of attention to utility is time-varying, expressing that early and later attention contribute differently to utility accumulation, which supports H3a. Second, the findings imply that both the trajectories of quantity and types of attention contribute to the accumulation of utility rather than only the trajectories of attention quantity or types, which supports H2a. Further hypothesis tests follow[5].

Contribution of Attention Quantity and Types. Table 2.4 gives parameter estimates of the full model, and of models 1 and 3 for comparison. Model 1 assumes that attention does not contribute to brand choice, and model 3 that only the trajectory of quantity of attention does. The contribution of attention to brand choice varies strongly across the time course of the choice task, as hypothesized. Specifically, the quantity of attention devoted to the brands (eye fixations) contributes stronger to brand choice in later than in earlier stages of the choice

[5] As robustness check, we tested whether the results of the proposed model are invariant to the experimental conditions and to the total decision time of participants. We added interaction terms between each of the twelve attention trajectory components, and respectively, the information density condition and the total eye fixation frequency for each consumer after grand mean-centering. The follow-up models each have 12 additional parameters but did not improve in fit over the proposed model.
We also calculated the hit rate for a model that predicts brand choice based on the brand fixed effects and brand ownership effects of the full model but restricts the contribution of attention to utility to zero. The hit rate of this model is 34%.

task. The 95% CI of the quadratic change in attention quantity [12.75; 37.42] does not overlap those of the initial level [-.18; 4.17] or linear effect [2.35; 9.97]. The estimate of the quadratic change of attention for integration (22.24, $p < .001$) is also larger than the estimates of the initial level (2.89, $p = .02$) and linear change (7.70, $p < .001$), although the 95% CIs overlap. Taken together this is further evidence that later attention quantity contributes stronger to brand choice than earlier attention does, in support of H3a.

Over and above the quantity of attention that brands receive, the contribution of attention for integration is statistically significant as well for the initial level (2.89, $p = .02$), linear (7.70, $p < .001$) and quadratic change (22.24, $p <.001$). Yet, only the initial level of attention for comparison is significant (3.16, $p = .01$), and none of the components of the trajectory of other attention are ($p > .14$). This supports H3b.

A comparison of the three models in Table 2.4 shows that brand ownership contributes to predicting brand choice when attention is not accounted for (model 1), reflecting a state dependence effect. Importantly, prior ownership does not contribute significantly anymore once attention to the brands is accounted for (model 3, $p =.09$; full model, $p = .18$), and the CIs with model 1 do not overlap. Moreover, Table 2.2 shows that brand ownership led to increased attention quantity (eye fixations) and attention for integration (within-brand saccades) for the owned brand. This reveals that the state dependence effect here is captured by the attention trajectories, rather than by mere choice inertia or habit formation. This supports H4.

A comparison of model 3 with the full model reveals that the contribution of the trajectory of attention quantity to brand choice drops when trajectories of attention types are added. Then, the contribution of the linear change in attention quantity more than halves from 14.13 ($p < .001$) in model 3 to 6.22 ($p < .001$), with the respective 95% CIs not overlapping. Likewise, the contribution of the quadratic change almost halves from 47.72 ($p < .001$) in

Table 2.4 Contribution of attention trajectories to brand choice

		Model 1				Model 3				Full Model			
Predictors		*M*	*p*-value	2.5%	97.5%	*M*	*p*-value	2.5%	97.5%	*M*	*p*-value	2.5%	97.5%
Brand 2	α_2	.24	.07	-.07	.57	-.36	.17	-.85	.19	-.74	.04	-1.45	-.04
Brand 3	α_3	-.70	.002	-1.15	-.22	-.49	.15	-1.21	.18	-.61	.11	-1.42	.14
Brand 4	α_4	.13	.18	-.21	.49	.36	.11	-.25	.99	.04	.25	-.79	.85
Brand 5	α_5	-.07	.59	-.43	.29	-.15	.53	-.78	.47	-.35	.37	-1.20	.51
Brand ownership	α_6	1.33	<.001	1.06	1.61	.29	.09	-.16	.71	.24	.18	-.41	.90
Attention Quantity:													
Initial level	β_0^1	--	--	--	--	4.17	<.001	3.43	4.98	1.96	.03	-.18	4.17
Linear change	β_1^1	--	--	--	--	14.13	<.001	12.22	16.27	6.22	<.001	2.35	9.97
Quadratic change	β_2^1	--	--	--	--	47.72	<.001	41.19	55.02	25.22	<.001	12.75	37.42
Attention Type:													
Integration:													
Initial level	β_0^2	--	--	--	--	--	--	--	--	2.89	.02	.17	5.74
Linear change	β_1^2	--	--	--	--	--	--	--	--	7.70	<.001	2.80	12.86
Quadratic change	β_2^2	--	--	--	--	--	--	--	--	22.24	<.001	10.94	33.35
Comparison:													
Initial level	β_0^3	--	--	--	--	--	--	--	--	3.16	.01	.41	5.98
Linear change	β_1^3	--	--	--	--	--	--	--	--	4.61	.08	-1.91	11.37
Quadratic change	β_2^3	--	--	--	--	--	--	--	--	4.62	.19	-8.81	18.33
Other:													
Initial level	β_0^4	--	--	--	--	--	--	--	--	-3.58	.23	-9.67	2.56
Linear change	β_1^4	--	--	--	--	--	--	--	--	-.55	.70	-9.61	8.69
Quadratic change	β_2^4	--	--	--	--	--	--	--	--	6.25	.14	-6.08	19.41

Note: M = Mean estimate; *p*-value = one-tailed Bayesian significance level; 2.5% and 97.5% are lower and upper bounds of credible intervals.

model 4 to 25.22 ($p < .001$) in the full model, with not overlapping 95% CIs. These results provide evidence that attention types, in particular attention for integration, contribute to brand choice over and above the mere attention quantity, and that eye saccades carry information about utility accumulation over and above fixations, supporting H2a and H2b.

2.5.3 Attention Trajectories towards Brand Choice

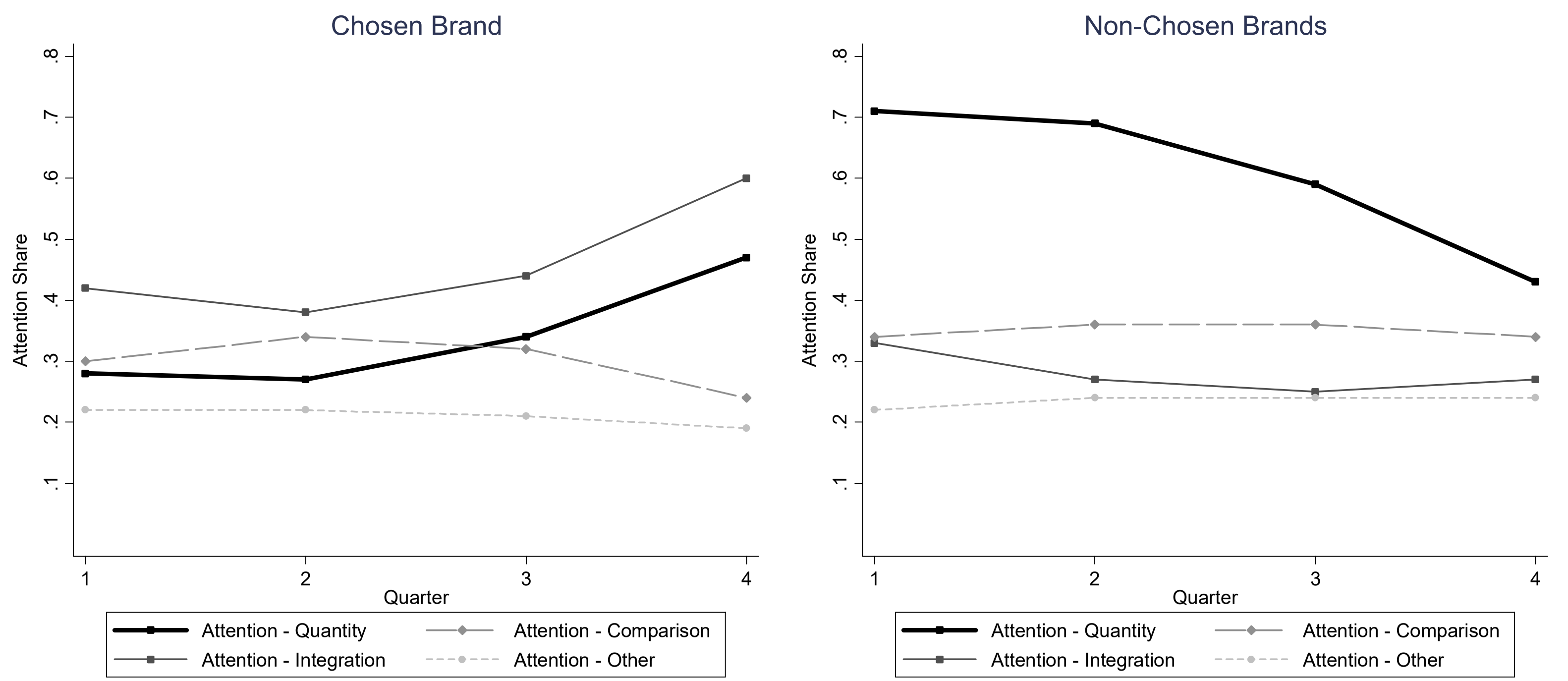
The analyses so far demonstrate that the contribution of attention to brand choice increased over time (β_k^g), but not whether attention to the chosen brand across the four time bins increased (θ_{ibk}^g). We estimated a slightly adapted version of the model in eqs. 3 and 4 to examine this. Specifically, eye-movement frequencies were converted into eye-movement shares and used as response variables (y) to facilitate comparison of heterogeneous consumers and eye-movement measures. Trajectories of attention shares were estimated by adding $\gamma_{k5}^g BC_{ib}$ to eq. 9, with the rest remaining unchanged. Here, BC_{ib} (1 = yes, 0 = no) indicates whether or not consumer i chooses brand b, and γ_{k5}^g captures differences between chosen and non-chosen brands in the attention share trajectories. Table 2.5 and Figure 2.3 document the attention trajectories for the chosen and non-chosen brands during choice.

There is evidence of a “double attention bias” towards the ultimately chosen brand: (1) it progressively attracts more of the total attention quantity during the choice task, and (2) a progressively larger share of its attention is allocated to information integration. First, with five brands in the choice set, 20% would be the “fair share” of attention quantity for the chosen brand which is close to the intercept of the initial level (.19, $p < .001$, Table 2.5, column 1). Yet, the chosen brand already attracts 8 percentage points more attention quantity after the first quarter of the task (.08, $p < .001$). The attention bias towards the ultimately chosen brand accelerates when choice implementation nears (quadratic change .04, $p < .001$). In the last quarter the chosen brand reaps an estimated 44% of the attention quantity, which is more than twice its fair share, with the four non-chosen brands jointly attracting the remaining 56%. Inspection of the raw data in Table 2.1 (final two columns) further illustrates this. Figure 2.3 plots the model-estimated shares of attention quantity for the chosen brand (thick line in left plot) and the non-chosen brands (thick line in right plot). Attention bias towards the ultimately chosen brand is not restricted to the final quarter but reveals itself early on and across the whole trajectory. This attention bias cannot be explained by multiple brands completely being ignored towards the end of the choice task, such that the chosen brand becomes the sole "attention survivor." Even in the last quarter, 60% of consumers still attended all five brands, 77% attended at least four, and 88% attended at least three.

Second, with three specific types of attention, 33% would be the fair share of attention for each type. This is the intercept of the initial level of attention for information integration (.33, $p < .001$). Yet, even after the first quarter the ultimately chosen brand already attracts 9 percentage points extra integration attention (.09, $p < .001$). This bias in integration attention for the ultimately chosen brand rises over time. The net effect of the linear (-.01, $p = .39$) and quadratic change (.03, $p < .001$) is positive. Figure 2.3 documents the resulting 18 percentage points increase in model-estimated share of integration attention for the chosen brand, from 42% in the first quarter to 60% in the last quarter. The shares of comparison and other attention for the chosen brand similarly drop over time. In sharp contrast, shares of the three attention types for the non-chosen brands remain essentially stable (intercepts of linear and quadratic change are not different from zero; Table 2.5). To illustrate, the share of integration attention for the non-chosen brand is 33% in the first quarter and 27% in the final quarter.

Taken together this supports the predictions from the proposed gSSM that both attention and its contribution to brand utility are time-varying. This is the first evidence that during choice the eventually chosen brand gains more of the total attention, that more of this attention is for information integration, and that the contribution of such attention to brand utility also grows.

Figure 2.3 Attention trajectories for chosen and non-chosen brands



Note: Estimated attention quantity shares sum to one for chosen and non-chosen brands combined. Parity share of attention quantity for the chosen brand is 20% per quarter (1 out of 5). Estimated integration, comparison, and other attention shares sum to one for chosen brand and non-chosen brands separately. Parity share is 33% for each attention type per quarter (1 out of 3).

Table 2.5 Attention trajectories towards brand choice

	Components of Attention Trajectories								
	Initial Level (k = 0)			Linear Change (k = 1)			Quadratic Change (k = 2)		
Predictors	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Attention Quantity									
Intercept	.19	.01	<.001	-.03	.02	.04	.001	.005	.43
Brand chosen	.08	.01	<.001	-.03	.01	.01	.04	.004	<.001
Brand owned	.02	.01	.05	.001	.02	.47	-.001	.01	.43
Heterogeneity (r^1_{ibk})	.02	.003	<.001	.01	.004	<.001	.001	.000	<.001
Attention Type									
Integration									
Intercept	.33	.02	<.001	-.08	.03	.01	.02	.01	.05
Brand chosen	.09	.02	<.001	-.01	.03	.39	.03	.01	<.001
Brand owned	.01	.02	.34	.01	.03	.35	-.01	.01	.28
Heterogeneity (r^2_{ibk})	.03	.01	<.001	.02	.01	<.001	.002	.001	<.001
Comparison									
Intercept	.34	.03	<.001	.03	.04	.22	-.01	.01	.28
Brand chosen	-.04	.02	.02	.04	.03	.10	-.02	.01	.03
Brand owned	.01	.02	.38	-.03	.03	.20	.01	.01	.20
Heterogeneity (r^3_{ibk})	.02	.01	<.001	.02	.01	<.001	.002	.001	<.001
Other									
Intercept	.22	.02	<.001	.02	.03	.29	-.005	.01	.31
Brand chosen	.004	.02	.40	-.02	.02	.17	.001	.01	.47
Brand owned	.01	.02	.22	-.01	.03	.41	.01	.01	.25
Heterogeneity (r^4_{ibk})	.01	.003	<.001	.01	.01	<.001	.001	<.001	<.001

Note: *M* = Mean estimate; *SD* = Standard deviation; *p*-value is one-tailed Bayesian significance level. Attention shares are examined. Integration, comparison, and other attention shares are relative to each other. Attention quantity shares are relative to other brands. Brand chosen: 1 = yes, 0 = no. Brand owned: 1 = yes, 0 = no. Effects of brand fixed-effects (brand-level), and of information density and product ownership (consumer-level) not shown to save space.

2.5.4 Model-based Inferences

The model allows us to infer the trajectories of utility and from these the choice probability trajectory of the chosen brand for each consumer using backward induction (eq. 7). That is, for the start of the choice task and for each of the four time periods during the choice task, we identify the brand most likely to be eventually chosen for each consumer, and compare this with the actually chosen brand. Choice probabilities at the start of the task (Q0), when no attention has been allocated yet, are based on model 1, and the rest on the full model. The accuracy of inferring final brand choice from the model estimates is termed “inference accuracy.” Because the inference can be accurate (1) or inaccurate (0) at each of the quarters

(Q1 to Q4) and before the start (Q0), there are 32 patterns of inference accuracy. Table 2.6 (column "all") and Figure 2.4 (thick line) summarize the results for all participants. The bottom part of Table 2.6 summarizes inference accuracy for each time period separately. Inference accuracy is already 56% after two quarters, reaching a high 77% after three quarters, and the final 85% at the end of the choice task. The thick line in Figure 2.4 shows these sharp increases in inference accuracy for the sample as a whole.

Expression and Accumulation of Utility. The specific accuracy patterns provide deeper insights. For a high 33% of the sample final brand choice could be accurately inferred in all four quarters during the choice task (patterns 31 and 32 in Table 2.6). These consumers where from start to end on-track to choose their preferred brand, attended to it most during the choice task and chose it. For a further 19% of the sample, brand choice could be accurately inferred in the final three quarters (patterns 29 and 30). These consumers quickly learned and "knew early" their preferred brand, attended to it most from there on and chose it. A further 21% of the sample "formed their final preference later" so that their final brand choice could be accurately inferred in the final two quarters (patterns 25 to 28). Thus, for a high 73% of the sample, final brand choice could be inferred at least in the final two quarters of the choice task. And for a low 9% of the sample (n = 32, patterns 17 and 18), final choice could be inferred only in the final quarter. These consumers apparently formed their preference for the finally chosen brand late in the task.

For another 4% of the sample (n = 13, patterns 9 to 16), final choice could be accurately inferred in the third but not the final quarter. These few consumers most likely engage in verification of other brands during the last quarter before implementing the final choice that they had made earlier (Russo and Leclerc 1994; Stüttgen et al. 2012).

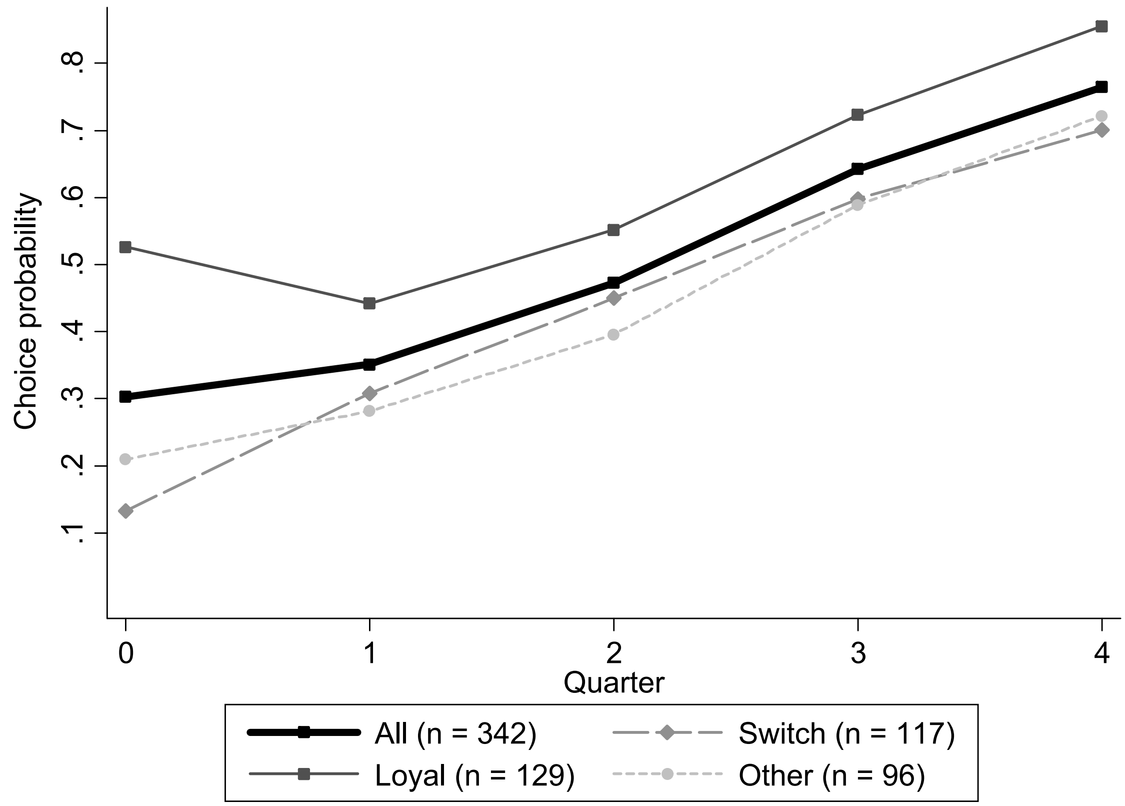
State Dependence Effects. The stratified sample of various brand owners and non-product owners also permits exploring state dependence effects on choice for complex

Table 2.6 Attention – choice link over time

	Inference Accuracy of Chosen Brand (1 = yes, 0 = no)					Consumer Segment			
Pattern	Start: Q0	Q1	Q2	Q3	End: Q4	All (n = 342)	Loyal (n = 129)	Switch (n = 117)	Other (n = 96)
1	0	0	0	0	0	21	0	15	6
2	1	0	0	0	0	9	6	0	3
3	0	1	0	0	0	4	0	4	0
4	1	1	0	0	0	1	1	0	0
5	0	0	1	0	0	0	0	0	0
6	1	0	1	0	0	1	1	0	0
7	0	1	1	0	0	2	0	0	2
8	1	1	1	0	0	1	0	0	1
9	0	0	0	1	0	5	0	1	4
10	1	0	0	1	0	1	1	0	0
11	0	1	0	1	0	0	0	0	0
12	1	1	0	1	0	0	0	0	0
13	0	0	1	1	0	4	0	4	0
14	1	0	1	1	0	1	0	0	1
15	0	1	1	1	0	1	0	0	1
16	1	1	1	1	0	1	0	0	1
17	0	0	0	0	1	23	0	12	11
18	1	0	0	0	1	9	8	0	1
19	0	1	0	0	1	2	0	1	1
20	1	1	0	0	1	2	2	0	0
21	0	0	1	0	1	0	0	0	0
22	1	0	1	0	1	0	0	0	0
23	0	1	1	0	1	0	0	0	0
24	1	1	1	0	1	4	3	0	1
25	0	0	0	1	1	43	0	21	22
26	1	0	0	1	1	26	21	0	5
27	0	1	0	1	1	1	0	1	0
28	1	1	0	1	1	3	2	0	1
29	0	0	1	1	1	38	0	22	16
30	1	0	1	1	1	26	26	0	0
31	0	1	1	1	1	46	0	36	10
32	1	1	1	1	1	67	58	0	9
			Accuracy (%) after:						
					Quarter 0	44	100	0	24
					Quarter 1	39	**51**	**36**	28
					Quarter 2	**56**	**68**	**53**	**44**
					Quarter 3	**77**	**84**	**73**	**73**
					Quarter 4	**85**	**93**	79	80

Note: Choice inferences at Q0 based on model 1. Loyal consumers choose the brand they currently own, hence 100% inference accuracy at Q0. Consumers who switch chose another than the currently owned brand, hence, 0% inference accuracy at Q0. Other consumers choose the highest choice share brand, which is 24% for brand 2. Bolded hit rates indicate significant difference ($p < .05$) from previous quarter for same segment. In all quarters, inference accuracy for loyals is higher than for switchers and others, which do not differ from each other, except after Q0, when inference accuracy for others is better.

Figure 2.4 Choice probability trajectory for the chosen brand



products. Note that choice probability of the loyal segment is 53% before starting the task and then drops in the first quarter to 44%. This drop is due to the initial attention that these consumers allocate to other brands than the brand they own and eventually choose again. Choice probability for the loyal segment increases to 55% after the first two quarters and eventually reaches 86%. Choice probability patterns for the other customer segments are similar but shifted down.

2.6 Discussion

This chapter set out to answer two key questions: (1) how do trajectories of attention to each of the brands during the choice task contribute to the accumulation of utility and final choice, and (2) which fundamental attention processes contribute to the accumulation of utility and brand choice. It was motivated by the RIT idea that, because attention is scarce, it is in consumers' best interest to bias their limited attention in favor of strong candidates and pay less attention to weak candidates in a choice task. Yet, strong and weak candidates are not always immediately evident to consumers and gradually emerge in particular in complex choice tasks. Attention trajectories document this accumulation of utility.

We proposed a generalized Sequential Sampling Model (gSSM) to answer these questions. It extends earlier models and research in three important ways. First, it decomposes overt eye-movement measures into covert attention and measurement error. This provides unbiased and dis-attenuated estimates of the link between eye movements, attention, and expected brand utility. Second, it separates the quantity of attention, as reflected in eye fixations, from the types of attention, as reflected in saccades between successive eye fixations. It identifies three types of attention, respectively, integration, comparison, and other. This provides insight into higher-order processes that may contribute to brand utility, and which cannot be readily identified otherwise. Third, it decomposes attention trajectories into key components: their initial level, and linear and quadratic change. This makes it possible to monitor when attentional preferences for brands surface during the choice task, and to quantify the contribution that the earlier and later stages have to utility and final choice. We estimated the gSSM on data from a large-scale eye-tracking experiment, with a national, stratified sample of 342 consumers. The results provide answers to both key questions, and empirical support for our hypotheses.

First, we found that covert attention was more intimately linked to brand choice than overt eye fixations were, in support of our hypotheses. Thus, prior research that has treated overt eye movements as mere replicates of covert attention may have systematically underestimated the link between attention and utility, and may have missed links where these exist.

Second, the trajectory of attention quantity contributed more to brand choice than the accumulated final sum of attention did. This supports the idea that the contribution of attention to utility is time-varying, and that later attention contributes more to utility than earlier attention.

Third, the trajectories of attention types contributed to brand choice over and above the trajectory of attention quantity. This supports the idea that eye saccades provide insight into utility accumulation over and above the sheer frequency of eye fixations. It also highlights the importance of controlling for attention type when quantifying the contribution of attention quantity, and that a failure to do so might overestimate the contribution of the latter.

Fourth, final brand choice revealed itself in attention trajectories early on. So it is not the case that attention only became biased towards the preferred brand just before expressing the final choice. In fact, the finally chosen brand already gained 8 percentage points more overall attention than its fair share after the first quarter of the task, and ended up with 24 percentage points more in the last quarter (44% of the total attention).

Fifth, in particular attention for information integration contributed significantly to brand choice, and in particular later during the choice task. This marks the importance of eye saccades within-brands which consumers deploy to form an overall evaluation of the merits and costs of a particular brand when faced with a complex choice, such as when choosing a smartphone. Thus, the finally chosen brand progressively gained more attention overall during the choice task, and an increasingly larger share of this attention was allocated to information integration rather than to brand comparison or other operations. Therefore, after half of the choice task, the final choice of 56% of the consumers already revealed itself in their attention patterns.

Sixth, state dependence effects, when consumers chose the new device of the brand they already owned, were fully captured by attentional trajectories to the brand, rather than being habitual or automatic. It was not the case that all loyal consumers only attended to the finally chosen brand; in fact they tended to allocate their early attention to other brands. Also, while information density influenced the duration of the choice task, it did not influence the attention trajectories nor the contribution of the attention trajectories to brand choice. This is suggestive of the fundamental nature of the link between trajectories of attention, the accumulation of utility, and final choice.

2.6.1 Implications and Future Directions

Our findings demonstrate the usefulness of the gSSM to examine (in)attention and its implications for choice. In support of the central tenet of RIT, consumers indeed did not allocate their attention proportionally across the brands, but instead biased their attention toward the brand they eventually chose, which has been documented before. Yet, a surprising result of our experiment is how early during the choice task the eventually chosen brand already attracted more than its fair share of attention. Thus, preferential allocation of attention was an emergent property that started much earlier than documented before. Prior research has reported a so-called "gaze cascade" where the finally chosen brand attracts a disproportionate amount of attention just before choice implementation. For instance, Shimojo et al. (2003) observed increased attention to the chosen alternative in the final second before choice. Shi et al. (2013) observed a similar effect in the final three fixations, which is about one second. Atalay et al. (2012) reported a gaze cascade effect in the final five seconds of choice for vitamins and food-replacement bars. If such an attention bias for the chosen brand would surface only late, it would be of theoretical but probably of lesser managerial and policy relevance. Our results are the first to document a much earlier attentional bias towards the ultimately chosen brand; revealing a “preferential attention trajectory” from the outset of brand choice in an information-rich environment. This opens up opportunities for novel theories and research about preference formation during choice tasks, and potentially managerial intervention if undesirable choice options appear to gain early traction.

The findings do not imply that consumers were completely path-dependent and universally chose the brand that initially garnered most attention. In fact, the findings reveal clear individual differences between consumers in how early the final winner revealed itself. Thus, 33% of consumers expressed preferential attention towards their final choice from the first quarter of the task onwards. Yet, 19% of consumers did so only after the first quarter, another 20% only after the second quarter, and 6% only after the third quarter. This reveals substantial variation in the accumulation of utility for the chosen brand.

An interesting avenue for further research and interventions is to assess the effects of intra-task changes in the information structure and content of the choice options. Such changes can occur exogenously, when a new brand or offer pop-up during the choice task that consumers are engaged in, or endogenously as a consequence of the attention trajectories themselves. For instance, contingent on the brand currently garnering preferential attention, new information about that brand or about competing brands might accelerate or decelerate utility accumulation. In a first stage, the gSSM could be trained on eye-movement and final choice data of participants to derive attention trajectories and brand utility weights. In a second stage, the model could then be used to predict final choice from the attentional trajectories of (other) participants before they implement their choice. We implement this approach in Chapter 3 in the context of consumer choice. The results help to understand and improve consumer choice, but perhaps can also be extended to medical and moral decision making (Pärnamets et al. 2015). In a recent review, Al-Moteri et al. (2017, p. 63) conclude that "…the investigation of eye-movement behavior in deliberate (analytical) decision-making modes does not appear to be a priority in eye-tracking studies in the medical field. This is an important area for future research." The proposed model and inference approach can inform such research and applications.

The present findings are the first, to our knowledge, to document the tight link between the accumulation of brand utility and trajectories of specific types of attention, in particular attention for information integration, over and above the mere quantity of attention. This extends RIT and prior SSMs which have emphasized the quantity of attention as the dominant or sole driver of utility. Yet, the same attention quantity to brands, as reflected in eye fixations, can be the result of qualitatively different specific types of attention, which may contribute differently to utility accumulation. Shi et al. (2013) examined the relationship between choice and the total duration that consumers were in a state of specific attention. Here, we quantify the contribution that the whole trajectory of different attention types for each brand has to utility, and over and above the trajectory of attention quantity, which is new. Our findings also suggest that choice models need to account for stickiness due to inattention to correctly model the impact of prior choice on current choices, because the processes accounting for state dependence (inattention, inertia, habit formation) have different managerial and public policy implications.

Our work has important limitations which point to future theory development and research. First, the findings support the idea that over and above the sheer amount of attention quantity, attention for information integration contributed to brand choice, but are contingent on the specific decision studied. Although the findings are broadly consistent with work that used information display boards to examine choice between gambles (Willemsen et al. 2011) and support RIT, we do not claim that our findings are general. The experiment concerned complex choice with a small choice and large feature set, and consumers who were in the market for the target product. When the choice set is large and the feature set is small, or when consumers are not yet in the market for a (new) purchase attention trajectories and their contribution to final choice might be different (Lohse and Johnson 1996; Shi et al. 2013). Future research with our model can test the role of choice and feature set sizes, and other contextual factors on the contribution that the various attention types have on choice.

Second, in an effort to keep the modeling tractable, to account for measurement error, and based on theory about stages in decision making, the eye-movement data were normalized into four time bins for each consumer. This is consistent with prior research that has used one up to four time bins (Meißner et al. 2016; Pieters and Warlop 1999; Willemsen et al. 2011), but limits the amount of detail about attention trajectories and utility accumulation. It also precludes modeling the time that consumers take to make a choice, which is an important caveat. Follow-up research which jointly models brand choice and individual decision time is particularly called for (Cavanagh et al. 2014). This is the focus of the model we develop in Chapter 3 of this dissertation.

Third, our model is agnostic about the causal processes linking attention and utility at each point in time, as other models are (Manohar and Husain 2013; Reutskaja et al. 2011). Thus, we cannot claim that attention causes utility, only that they are empirically strongly associated. The observed systematic links between trajectories of attention, utility and choice, informed by theory, do suggest such a causal link, but our data and model do not permit strong causal inferences. It is reasonable to assume that attention and utility are part of a positive feedback loop (Shimojo et al. 2003), with utility driving attention (consumers search for value; Yang et al. (2015)), and attention driving utility (via changes in the utility weights; Bordalo et al. (2013)), with reasonable time-out mechanisms. Whereas both causal directions have been reported, we are not aware of research that directly examines and quantifies the moment-to-moment bidirectional effects between attention and utility, which is an important direction for future research.

Linking work on attention and choice to the literature on sequential search for information and choice would prove very useful in establishing such structural and causal links. Especially work on repeated search decisions leading to a (single) purchase decision would be relevant. There is initial work to formalize costly search for information before obtaining the rewards from a choice between multiple products (Ke, Shen, and Villas-Boas 2016) or a choice of a single product with multiple features (Branco, Sun, and Villas-Boas 2016). This type of work relies on various simplifying assumptions and as the number of attributes and options rise, estimating such models becomes intractable. Integrating our descriptive attention-and-choice model with search-and-choice theories is challenging but a potentially fruitful area for cross-disciplinary work, novel theories and empirical findings. Such future work will add to the current effort to identify fundamental links between eye movements, attention, and utility, and to gain deeper insights into value-based decision processes.

Chapter 3

How Attention Reveals Why Consumers Choose What When

3.1 Introduction

Consumers use online environments to make choices, such as selecting health insurance or making a purchase for a durable good. This facilitates access to detailed information that consumers could inspect in order to identify the choice alternative that provides the best utility match. Yet, consumers (1) acquire only some information about a limited set of alternatives. Even more, by the time they make a brand choice, consumers (2) will have repeatedly inspected information about the brand that they eventually choose and (3) will have done so increasingly towards the end of the decision process. These three characteristics of how consumers inspect brands before choice have been documented with observational data that covers weeks of online search (Bronnenberg et al. 2016) and in controlled studies that facilitate information acquisition by presenting brands side-by-side (Meißner et al. 2016; Shi et al. 2013; Yang et al. 2015).

Models in the bounded rationality literature (Simon 1955) found that information acquisition and processing costs motivate consumers to inspect only parts of the available information while leaving others unexplored. While this justifies the first of the three characteristics of how consumers inspect brands before choice, the explanation for the other two remains unknown. Namely, it is not clear why consumers are more likely to repeatedly inspect information about the brand that they eventually choose and why they do so increasingly towards the end of the decision process. These cannot be explained by information processing costs as there is no compelling reason why consumers would be more likely to forget information about the brand that they eventually choose closer to the moment when they express their choice.

Building on rational inattention theory (RIT) (Caplin and Dean 2015; Steiner et al. 2017) and sequential sampling models (SSM) (Krajbich et al. 2010; Ratcliff and Smith 2004), we propose that consumers decide from moment-to-moment (1) what brand to inspect and (2) for how long in order to reduce uncertainty about the utilities of the brands on display. This does not exclude the impact of information processing costs, but suggest that consumers are more likely to verify information for brands that are worth the effort. Therefore, differences in attention to brands are closely linked to brand utilities and choice probabilities. The sequential nature of eye movements that consumers use to inspect the brands offers the opportunity to predict both brand choice and its timing from moment to moment, as the consumer progresses through the task.

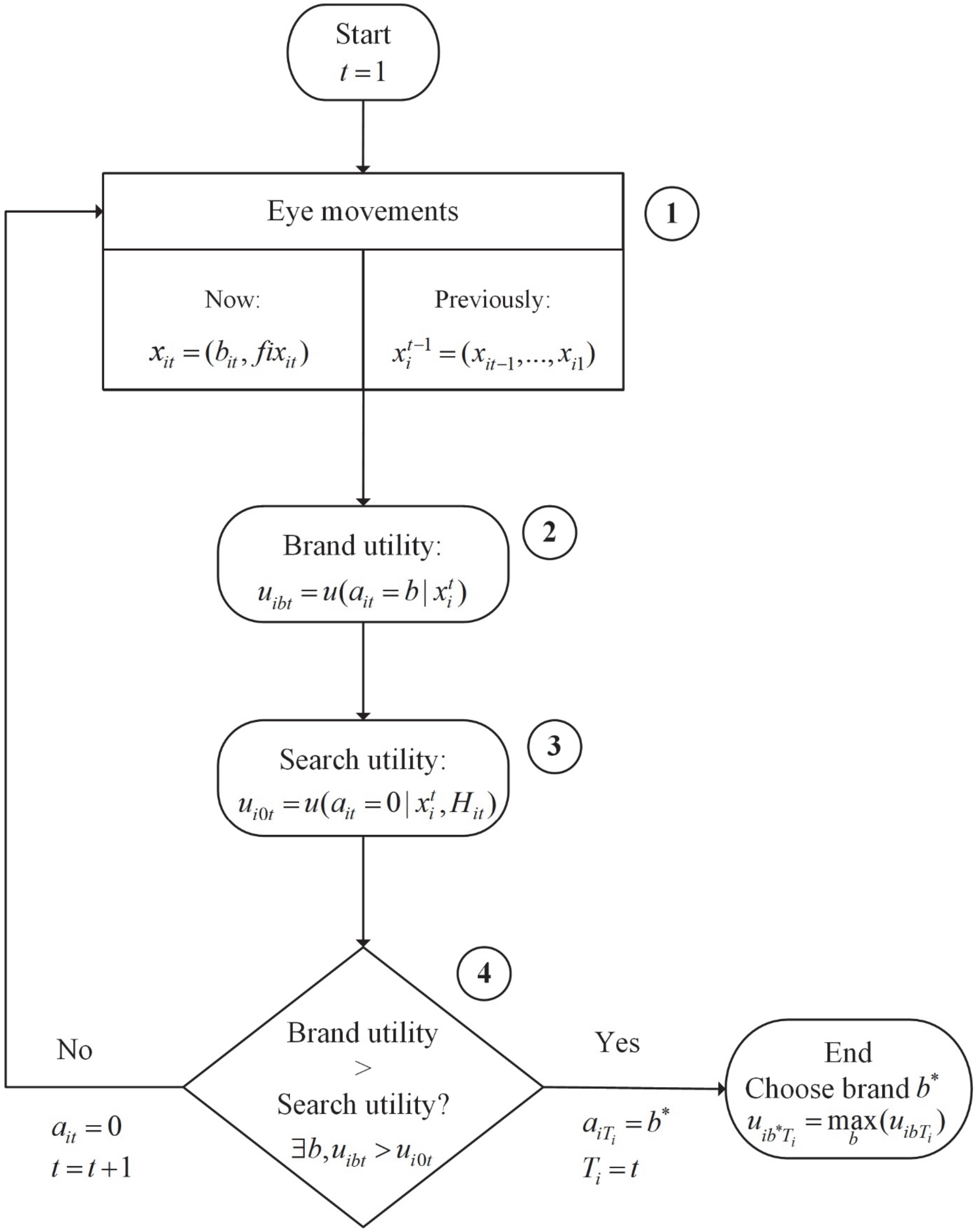
In this chapter we develop a model that provides insights into how consumers decide (1) what brand to choose and (2) the moment when to express their choice, and (3) how these decisions are linked to the sequential eye movements that precede them. The model specifies that consumers use eye movements to inspect the brands on display until choosing one of them justifies foregoing the net benefits of additional information. The model is calibrated on eye-tracking and choice data from a study ($N = 214$) on single choice between four digital cameras for which information is displayed side-by-side. We use eye movements that consumers make during the task, as manifest indicators of information acquisition and attentional processes (Just and Carpenter 1980; Shi et al. 2013), to gain insights into rapid and partly automatic moment-to-moment decision-making processes (Lohse and Johnson 1996; Russo and Leclerc 1994; Wedel et al. 2008). We use the results of the proposed model to test the empirical support for implications of SSM and RIT. The rest of the chapter is organized as follows. In the next section we present the model and explain how it relates to existing theories. Then, we describe the study, the econometric specification of our model, and the results. We end with discussion and implications.

3.2 Eye movements, Utility Accumulation, and Brand Choice

We propose an attention and utility accumulation model to formalize the link between eye movements and brand choice during a single consumer decision. Our approach focuses on choice from a set of B brands for which information is displayed side-by-side, such as on comparison websites. In this setting, consumers inspect the brands by moving their eyes to different areas of the display. During this visual inspection process, the eyes move rapidly from one location of the display to another. Eye-fixations are brief moments (200-400 msec.) when the consumer acquires information from a specific location that the eye is fixated on (Rayner 1998). In what follows, we index brands with b ($b = 1, \dots, B$) and moments during the decision process with t ($t = 1, \dots, T_i$), where B is the number of brands on display and T_i is the moment when brand choice is expressed. One moment corresponds to a sequence of two or more consecutive fixations on the same brand. For example, if a consumer uses four consecutive eye-fixations to inspect brand 1, these correspond to a moment t when b = 1 is inspected using fix_{it} = 4 eye-fixations (section 3.3.3 provides a more detailed description).

Consumers are strategic about what brand they inspect and for how long they do so from moment-to-moment before choice. Specifically, they inspect in more detail brands that appear promising and pay minimal or even no attention to those that are not worth the effort. This implies that sequences of eye-fixations are closely linked to the moment-to-moment accumulation of utility that takes place until the final brand choice is observed. The model has four components: eye movements (eq. 1), brand utility (eq. 2), search utility (eq. 3), and moment-to-moment utility comparison (eq. 4). Figure 3.1 offers a visual representation of the model and the four components. In Table 3.1 we summarize the model components and assumptions, their implications, and the empirical evidence needed to support them. We now provide a concise description of the decision process for a consumer i and in the next section discuss how the model components and their determinants are related to RIT and SSM.

Figure 3.1 Eye movements, utility accumulation, and brand choice



Note: $x_{it} = (b_{it}, fix_{it})$ indicates that at moment t consumer i uses fix_{it} eye fixations to examine brand b_{it}; $b_{it} = 1, \ldots, B$; $x_i^t = (x_{it}, \ldots, x_{i1})$ is the sequence in which consumer i inspects the brands on display up to moment t; $b = 1, \ldots, B$; H_{it} is the uncertainty experienced by consumer i at moment t; b^* is the brand with maximum utility among the B brands on display; T_i is the moment of choice for consumer i; a_{it} is the action consumer i takes at t.

At each moment t ($t = 1, \dots, T_i$) during a choice task, the consumer uses eye-fixations to inspect the brands on display. We use $x_{it} = (b_{it}, fix_{it})$ to indicate the brand (b_{it}) inspected by the consumer i at time t using fix_{it} eye-fixations. Let $x_i^t = (x_{it}, \dots, x_{i1})$ be the time-sequence in which the consumer examines the brands on display up to moment t. For example, if at $t = 3$ the consumer has two fixations on brand 4 ($x_{i3} = (4, 2)$) and before that had three fixations on brand 3 ($x_{i2} = (3, 3)$) and four on brand 1 ($x_{i1} = (1, 4)$), then $x_i^3 = (x_{i3}, x_{i2}, x_{i1}) = ((4, 2), (3, 3), (1, 4))$.

The consumer uses the information acquired up to moment t to calculate two types of utility: (1) brand utility and (2) search utility derived from inspecting the brands longer. Then, the consumer compares these utilities and takes one of two possible actions: (1) end the decision process by choosing the brand that provides the largest utility ($a_{it} = b^*$, where $\max_b u_{ibt} = u_{ib^*t}$) and (2) continue to inspect the brands ($a_{it} = 0$). The consumer expresses brand choice when the utility of choosing a brand is larger than that of continuing the search.

Eye movements: $$x_{it} = f_{it}(u_i^{t-1}, x_i^{t-1}) + \xi_{it} \tag{1}$$

Brand utility: $$u_{ibt} = u(a_{it} = b | x_i^t) \tag{2}$$

Search utility: $$u_{i0t} = u(a_{it} = 0 | x_i^t, H_{it}) \tag{3}$$

Utility comparison:

$$a_{it} = \begin{cases} b^*, \ if \ \max_b u_{ibt} \geq u_{i0t} \ and \ \max_b u_{ibt} = u_{ib^*t} \\ 0, \ if \ \max_b u_{ibt} \leq u_{i0t} \end{cases} \tag{4}$$

Where: $x_{it} = (b_{it}, fix_{it})$ indicates the brand b_{it} examined by consumer i at moment t using fix_{it} eye-fixations; $b_{it} = 1, \dots, B$; f_{it} is the information acquisition strategy that consumer i uses at moment t (described in section 3.2.1); ξ_{it} is an unobserved random error (described in section 3.2.1); a_{it} is the action consumer i takes at moment t; $x_i^t = (x_{it}, \dots, x_{i1})$ is the sequence in which consumer i examines the brands on display up to moment t; $b = 1, \dots, B$; H_{it} is the uncertainty at moment t for consumer i (explained in section 3.2.2); b^* is the brand with maximum utility among the B brands on display; $t = 1, \ldots, T_i$ where T_i is the moment of choice for consumer i. Subscript t indicates values at moment t while superscript t indicates a vector of length t that contains all values up to and including moment t. For example, $a_i^t = (a_{it}, \ldots, a_{i1})$ is the sequence of actions of consumer i up to and including t.

3.2.1 Eye Movements

Costly information acquisition motivates consumers to be strategic in how they inspect the brands and what information they attend to (Steiner et al. 2017; Yang et al. 2015). After an initial orientation stage during which the brands are only briefly looked at, consumers inspect in more detail those brands that appear promising (Russo and Leclerc 1994). Hence, the decision of what brand and for how long to inspect depends on the current evaluation of the brands and the information previously acquired (Stüttgen et al. 2012; Willemsen et al. 2011; Yang et al. 2015). To the extent that (1) consumers inspect the brands they are more likely to choose and (2) consumers prefer different brands, there is brand and consumer heterogeneity in eye-movement sequences.

Eq. 1 formalizes three important aspects of the link between eye movements and utility. First, the model specifies that consumers decide which brand (b_{it}) and for how long (fix_{it}) to inspect based on what they previously fixated on (x_i^{t-1}) and the corresponding accumulation of utility (u_i^{t-1}). This is consistent with evidence that consumers revisit brands that they eventually choose (Bronnenberg et al. 2016; Krajbich et al. 2010), and suggests that the amount and timing of attention reflect how the consumer evaluates the brands. This behavior has been documented even when consumers are expected to use a strategy that renders revisits improbable, such as a satisficing strategy (Stüttgen et al. 2012). Second, the link between eye movements and utility is consumer and time specific (f_{it}). This accounts for evidence that consumers use various decision strategies that lead to different amounts and types of information being acquired (Bettman et al. 1998) and that they frequently switch between such strategies during the process (Shi et al. 2013). Third, the model accounts for unobserved error (ξ_{it}) in the mapping between what consumers intend to focus on and the exact location of their eye-fixations, which can be due to salience (Hutton 2008; van der Lans et al. 2008b) and corrective eye movements (Reichle and Drieghe 2015).

3.2.2 Utility Accumulation

Rational Inattention Theory asserts that "limits on attention impact choice" (Caplin and Dean 2015, p. 2183). Recently, analytical models of optimal information-processing behavior have been proposed (Matějka and McKay 2015; Steiner et al. 2017) that as far as we know have not been tested empirically. We build on these theoretical developments and offer a model that is calibrated on eye movements and brand choice data. In this section we describe how the two types of utility (brand and search) are related to the attention reflected by eye-movement sequences.

Brand utility. The amount of attention that the consumer allocates to each of the brands is reflected in the total number of eye-fixations they receive, which is linked to choice probabilities and preference (Glaholt et al. 2009; Krajbich et al. 2010; Pieters and Warlop 1999). When the consumer finds a brand interesting enough to inspect in more detail and use in comparisons with other brands, the number of fixations on this brand increases over time. This reflects a larger amount of attention on the brand ($\bar{x}_{ibt}$) relative to the other brands on display. While over time consumers can inspect two brands for about the same time, at a specific moment t their eyes can only fixate on one of the brands. The moments when consumers inspect a brand reflect how their interest in each of the brands grows or diminishes as they come across new information or as they become more certain of the brand utilities. More specifically, brands that have recently been fixated on are more likely to be considered as compared to brands that have not been inspected in a while, even if the consumer examined both of them for a similar amount of time. This implies that brand utilities are closely linked both to the extent that the brand has been under focus ($\bar{x}_{ibt}$) and to the moments when this took place ($\bar{\bar{x}}_{ibt}$).

Search Utility. Consumers balance the costs and benefits of examining the brands longer when they decide whether to choose one of the brands at current time t or to postpone the choice for a later moment $T_i > t$ (Krajbich et al. 2010; Meißner et al. 2016; Yang et al. 2015). Before acquiring any information about the brands, consumers are uncertain about which of the brands offers the best utility match. By examining the brands on display, consumers reduce uncertainty about the brand utilities and are better able to make a choice. As the brands are inspected in more detail, the utility of continuing the search reduces unless consumers acquire information that makes them less confident about their final choice. This implies that consumers derive more utility for search when they are uncertain (high H_{it}) which of the brands is a better match than when they are confident about which of the brands is best (low H_{it}). This is consistent with evidence that consumers spend more time in the decision process when there are two or more good alternatives than when one of them dominates the others (Krajbich et al. 2010; Meißner et al. 2016)

3.2.3 Utility Comparison

The consumer ends the decision process when the utility of one of the brands (u_{ib^*t}) is larger than that of additional search (u_{i0t}). Therefore, u_{i0t} is a decision threshold that one of the brands needs to reach in order for choice to be expressed. Our model accounts for three determinants of the moment of choice: (1) the utility that the consumer derives by ending the search and choosing the best brand, (2) uncertainty, and (3) time costs, as we explain in more detail in section 3.4 (Econometric Specification).

3.2.4 Contribution

Standard rational choice models make the assumption that consumers have all information about the brands and choose the brand with maximum utility. However, consumers rarely

Table 3.1 Attention and utility accumulation model – summary of the main components, assumptions, implications, and tests

	Model Component	Assumption	Implication	Test
1	*Eye movements* are used by consumers to inspect the brands on display. $x_{it} = (b_{it}, fix_{it})$ (eq. 1)	Consumers are strategic in choosing what brands, and for how long, they examine at each moment during the decision process. The brand that a consumer attends to at a specific moment is relevant at that time. Consumers are more likely to examine repeatedly better choice candidates than brands that they no longer consider.	There is consumer and brand variability in attention. Consumers allocate more attention to promising brands. Consumers no longer acquire information about brands with lower choice probabilities.	$\Sigma_{10} \neq 0$ (eq. 7) $\Sigma_{01} \neq 0$ (eq. 8) $\beta_1 > 0$ (eq. 9) $\beta_2 < 0$ and $\beta_3 > 0$ (eq. 9) $PP_6^{BC} > PP_2^{BC}$
2	*Brand utility* is the utility that the consumer derives by terminating search and choosing a brand. $u_{ibt} = u(a_{it} = b \mid x_i^t)$ (eq. 2)	After consumers identify a promising brand, the utility of that brand increases over time until it reaches a threshold and brand choice is observed.	Brand choice can be predicted earlier than it is observed with increasing accuracy as more eye movements are observed.	PP_6^{BC} increases over time
3	*Search utility* is the utility that the consumer derives by continuing to acquire information about the brands. $u_{i0t} = u(a_{it} = 0 \mid x_i^t, H_{it})$ (eq. 3)	Uncertainty (H_{it}) is the extent to which the consumer is unsure which of the brands offers the most utility. Consumers who are uncertain benefit from acquiring additional information.	Uncertainty has a positive effect on search utility. The moment when consumers express brand choice can be predicted with increasing accuracy as more eye movements are observed.	$\gamma_1 > 0$ (eq. 12) $PP_6^{MC} > PP_5^{MC}$ PP_6^{MC} increases over time
4	*Moment-to-moment utility comparison* is the process whose outcome is the consumer action (a_{it}) at time t. u_{i0t} vs. u_{ibt} (eq. 4)	At each moment before brand choice is expressed, consumers compare the utility they derive from choosing a brand to the utility of continuing the search and then choose the action with highest utility.	Search utility is a decision threshold that brand utilities need to reach in order to be chosen. Consumers forego the benefit of reducing uncertainty when the utility they derive from choosing a brand is larger.	$PP_6^{MC} > PP_4^{MC}$

Note: $x_{it} = (b_{it}, fix_{it})$ indicates that at moment t consumer i uses fix_{it} eye-fixations to examine brand b_{it}; $b_{it} = 1, \dots, B$; a_{it} is the action consumer i takes at t; $x_i^t = (x_{it}, \dots, x_{i1})$ is the sequence in which consumer i acquires information about the brands on display up to moment t; PP_m^y is the predictive performance of model m for outcome y; $m = 6$ for the proposed model; $m = 1, \dots, 5$ for models M1-5 in Table 3.2; $y = BC$ (brand choice), and MC (moment of choice).

acquire all available information, which impacts their brand choice. This has motivated the development of new theories and models, such as rational inattention theory (RIT) (Matějka and McKay 2015; Steiner et al. 2017) and sequential sampling models (SSM) (Krajbich et al. 2010; Ratcliff and Smith 2004), that explain choice while accounting for limited information acquisition and processing. The model developed in this chapter builds on these theories. It jointly models the two outcomes of the decision process (the chosen brand and the moment of choice) and contributes to the literature in three important ways.

First, the model predicts what consumers are going to choose from moment-to-moment as they inspect the brands before expressing their choice. We do this by using differences in the amount and timing of attention that each brand receives to infer how (1) consumers evaluate the brands at every moment before choice is observed and how (2) these evaluations change as consumers continue to inspect the brands. This differs from previous models that provide insight into how consumers evaluate the brands only after brand choice is observed, but that cannot be used to predict choice during the process. Such models focus on differences between brands over predefined time intervals relative to the observed moment of choice: four equal time bins (previous chapter), the two halves of the decision process (Meißner et al. 2016), or the final few seconds before choice (Atalay et al. 2012; Shimojo et al. 2003).

Second, the model jointly models the final brand choice and the moment when it is expressed, which has been indicated as an important research area (Cavanagh et al. 2014; Krajbich, Oud, and Fehr 2014). The model predicts both the eventually chosen brand and the moment of choice for new consumers and updates these predictions from moment-to-moment as more eye-fixations are observed. This is not possible using previous methods as they require prior preference measurements (Krajbich et al. 2010; Reutskaja et al. 2011), previously observed repeated choices (Meißner et al. 2016; Yang et al. 2015), or the complete sequence of eye movements (Atalay et al. 2012; Shimojo et al. 2003). While our empirical application makes use of data from a controlled eye tracking experiment, the model is more widely applicable, as we explain in the general discussion section.

Third, the model investigates how consumers balance the costs and benefits of inspecting the brands longer and quantifies their effects on decision time. This has important implications for understanding consumer variation in decision time and extends previous SSMs, such as the aDDM, which assume a consumer-invariant threshold fixed prior to the decision task (Ashby et al. 2016; Krajbich et al. 2010). Understanding how consumers adjust the decision threshold from moment-to-moment during preference-based choice provides an important extension of previous applications of SSM in perceptual decision making (Cavanagh et al. 2014; Forstmann et al. 2016).

3.2.5 Model Comparison

We test the proposed model by comparing it against two groups of competing models (Table 3.2). The first group contains models M1 and M2 that have different assumptions about the link between eye movements and brand utility. The second group contains three models (M3, M4, and M5) that specify different determinants of the moment when brand choice is expressed. Comparing our model (M6) against these specifications test assumptions about the link between eye movements and brand choice, and sources of variation in decision time.

M1 specifies that the mapping from brand utilities to eye-fixations occurs without error ($\xi_{it} = 0$, eq. 1), which is a common assumption in previous work (Cavanagh et al. 2014; Krajbich et al. 2010; Pieters and Warlop 1999; Reutskaja et al. 2011). Chapter 2 reports initial evidence against this assumption. Furthermore, this chapter investigates the impact of accounting for such unobserved error for predicting the moment when brand choice is expressed.

Table 3.2 Predictive performance at the moment of choice

	What brand to choose			*When* to choose				Out-of-sample performance at T_i	
	Attention			Moment-to-moment comparison of search utility against		Uncertainty	Cost of time	Brand Choice (*what*)	Moment of choice (*when*)
	Amount		Timing	Fixed decision threshold	Brand utilities				
	$\bar{x}_{ibt}$	y_{ibt}	$\bar{\bar{x}}_{ibt}$	$u_{i0t} > 0$	$u_{i0t} > \max_b u_{ibt}$	H_{it}	$\gamma_2 t$	PP_M^{BC}	PP_M^{MC}
M1		x			x	x	x	57.2%	53.5%
M2	x				x	x	x	54.5%	72.7%
M3	x		x	x			x	59.5%	62.3%
M4	x		x	x		x	x	59.7%	67.6%
M5	x		x		x		x	59.6%	73.3%
M6	x		x		x	x	x	59.6%	73.4%

Note: all models include brand intercepts (α_b).

Similar to our proposed model, M2 accounts for imperfect mapping between what utility and eye-fixations (ξ_{it}, eq. 1). However, M2 assumes that the moments when the consumer examines each brand does not reflect brand utilities over and above the amount of attention. This is a core assumption of analytical models of RIT (Matějka and McKay 2015; Steiner et al. 2017).

M3 specifies that consumers inspect the brands as long as continuing the search provides them with non-negative utility. This assumption is consistent with applications of the aDDM in which consumers use a fixed decision threshold set before the start of the process. If M3 provides a better fit than the proposed model, then the decision to continue the search is not influenced by consumers' uncertainty (H_{it}) or by the brand utility consumers would derive by expressing choice ($a_{it} = 0\ if\ u_{i0t} > 0$ and $a_{it} = b^*\ if\ u_{i0t} < 0$, where $\max_b u_{ibt} = u_{ib^*t}$, eq. 4).

M4 relaxes M3 by specifying that the extent to which the consumer is uncertain about which brand offers more utility (H_{it}) influences search utility. This is in line with previous applications that have found longer decision durations when choice alternatives provide very similar utility (Krajbich et al. 2010; Krajbich and Rangel 2011). Same as M3, M4 assumes consumers' decision to continue inspecting the brands is not influenced by the utility of the best brand on the display.

Different from M3 and M4, M5 assumes that consumers compare the utility they can derive by choosing the best brand against the utility of continuing the search. However, M5 specifies that the decision to continue the search is not influenced by uncertainty. This implies that consumers do not benefit from additional information that can help them differentiate between the brands as long as choosing any of them provides sufficient utility. While this assumption is typical for DDM, it is not always supported by applications that include eye-movement data, as results suggest consumers spend more time choosing between two or more good alternatives than when one of the brands dominates the others (Cavanagh et al. 2014; Krajbich et al. 2010; Reutskaja et al. 2011).

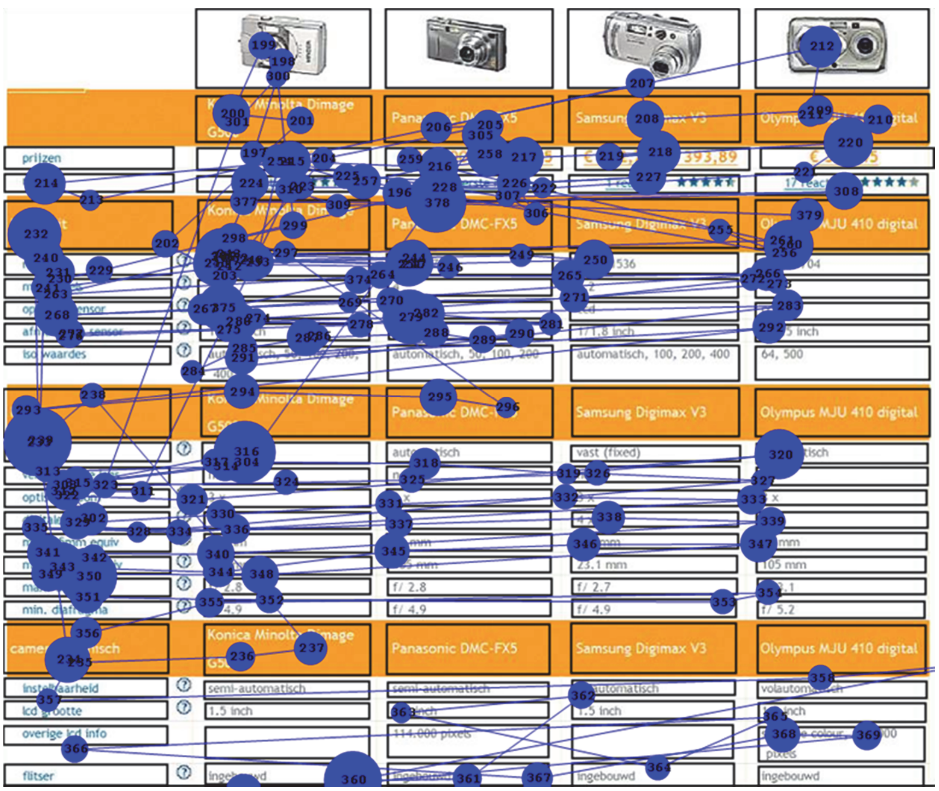
Empirical support for our model would imply that (1) both the amount of attention and the moments when the brands attracted it reflect brand utilities, (2) it is necessary to account for unobserved error in the mapping between utility and eye-fixations, (3) consumers express brand choice when brand utility reaches a time and consumer specific decision threshold, and (4) uncertainty motivates consumers to inspect the brands longer and postpone brand choice.

3.3 Data

3.3.1 Participants and Design

Students at a large public university ($N = 214$) participated in a study on consumers' choice for a digital camera. The experiment lasted approximately ten minutes and it was part of a one-hour session with other unrelated studies. Participants received the following instruction: "You have decided to purchase a digital camera today. Your price range is €300-400 and you are buying the digital camera online. Afterwards you will carefully examine your expectations and the experiences with the camera." The experiment had a between-subjects design and participants were randomly assigned to one of the following five conditions: (1) avoiding regret about the outcomes (40 participants), (2) avoiding regret about the process (41 participants), (3) avoiding disappointment (43 participants), (4) following the brain (rational, cognitive) (41 participants), and (5) following the heart (affective, emotional) (49 participants). These conditions mimic different decision goals that are expected to influence how participants inspect the brands (Bettman et al. 1998; Shi et al. 2013). After reading specific instructions depending on the condition they were assigned to, participants were shown the attribute-by-brands matrix (Figure 3.2). After choosing one of the cameras, participants received feedback about their choice (on screen) and answered a questionnaire about their expectations and experience with the camera and the choice. The last slide had a message to thank the participants for their cooperation. All participants were paid 7$ for participating in the one-hour session. Responses to the post-choice questionnaire are not used in the current study, and the goal manipulation had minimal effects on attention and choice and is not further considered.

Figure 3.2 Stimulus and example of gaze plot for one participant



3.3.2 Eye-tracking Procedure

Participants were informed that their eye movements were going to be recorded while they made a choice between four digital cameras. The experimental room was a cubicle containing a Tobii 1750 eye-tracker arranged on a table and a chair in front of it. Participants were seated in front of the screen such that the center of the screen was on the same level as their eyes. The distance between the eyes and the screen was 60 centimeters and all tracking was binocular at a frame-rate of 50Hz. Participants saw the instructions and stimuli slides projected on the screen and could move to the next slide by pressing the space bar. After deciding which brand to choose, participants could indicate their choice by fixating on the chosen brand and pressing the space bar to record their choice. Slide projectors, eye recordings, and brand choice recordings were computer controlled.

3.3.3 Grouping Fixations into Moments

A fixation is defined as a relatively stable eye-in-head position within some threshold of dispersion (usually ~2°) and with a velocity of 15-100 degrees per second (Duffy, Morris, and Rayner 1988). This study uses number of fixations on specific relevant areas of the stimulus display, similar to other research on eye movements (Pieters and Warlop 1999).

Eye-tracking data offer a detailed account of the areas attended to during the choice process. However, a single fixation is not enough to process detailed textual or numerical product details (Glaholt, Wu, and Reingold 2010; Stüttgen et al. 2012; Yang et al. 2015). Eye fixations are grouped into more meaningful measures of brand focus named moments, similar to previously work (Meißner et al. 2016). One moment corresponds to a sequence of two or more consecutive fixations on a brand, separated by at most one fixation on a different brand. For example, the fixations in the sequence ‘b1 b1 b1 b4 b1 b4 b3’ correspond to two moments. During the first moment ($x_{i1} = (1, 4)$) the consumer uses four eye-fixations to inspect brand 1 and during the second moment ($x_{i2} = (4, 2)$) two fixations are used to inspect brand 4. Let fix_{ibt} be the number of fixations on brand b at moment t:

$$fix_{ibt} = \begin{cases} fix_{it}, & if\ the\ consumer\ fixates\ on\ brand\ b \\ 0, & otherwise \end{cases} \tag{5}$$

The total number of fixations that brand b receives up to and including moment t is $cfix_{ibt} = \sum_{v=1}^{t} fix_{ibv}$ and $cfix_{ib}^{t} = (cfix_{ib1}, \dots, cfix_{ibt})$ is the sequence of fixations cumulated by brand b. There are B such sequences, one for each brand on display, and they indicate the timing and amount of attention allocated to each of the brands thus far.

Brand choice shares are 42% (brand 1), 12% (brand 2), 7% (brand 3), and 39% (brand 4) (Figure 3.3). Participants in our study express their final choice, on average, after 21 brand visits regardless of the brand that they eventually choose (*F*-test = 1.48, *p*-value = .221) (Figure 3.4). While the chosen brand receives more than 25% of fixations (*M* = 35%, *p*-value <.001) and visits (*M* = 32%, *p*-value <.001), all of the brands on display are inspected at least once by 91% of the participants.

Figure 3.3 Number of participants who choose one of the brands

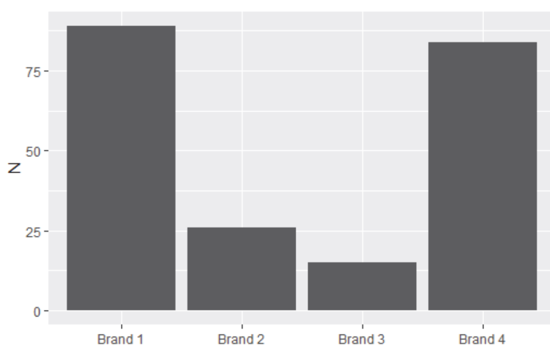
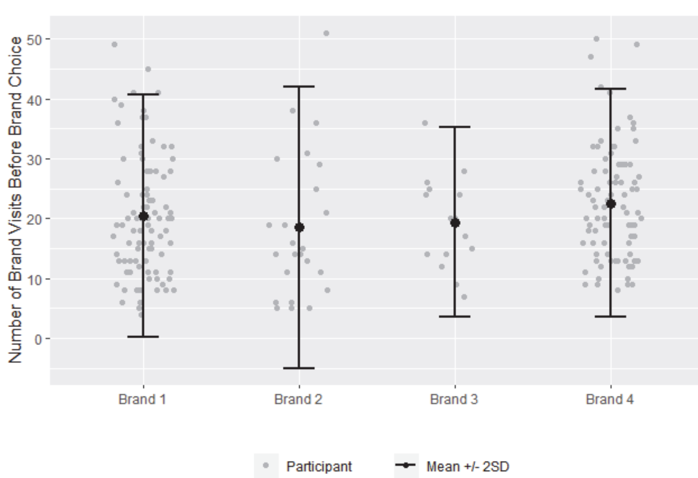


Figure 3.4 Participants spend similar amounts of time regardless of what brand they choose



3.4 Econometric Specification

3.4.1 Eye Movements

We account for consumer specific decision strategies and brand evaluations that impact the number, frequency, and pattern of fixations (Bettman et al. 1998; Shi et al. 2013; Stüttgen et al. 2012) by decomposing the observed eye-fixations into a consumer-specific level of attention ($\bar{x}_{i0t}$), brand-specific deviations ($\bar{x}_{ibt}$) from this average, and measurement error, similar to the approach in Chapter 2:

$$y_{ibt} = \bar{x}_{i0t} + \bar{x}_{ibt} + \xi_{ibt} \tag{6}$$

where ξ_{ibt} is unobserved heterogeneity and $y_{ibt} = \ln(cfix_{ibt} + 1)$ is the log transformed cumulated number of fixations. We use this log normal transformation to normalize the distribution of fixations (Pieters and Wedel 2004).

The model accounts for consumer-level differences in attention due to the decision goal conditions that participants are assigned to (explained in section 3.3.1) and due to unobserved heterogeneity as well as brand-level differences in attention due to the order in which the brands are display and due to unobserved heterogeneity:

$$\bar{x}_{i0t} = \left(\eta_{00} + \eta_{g0} + \epsilon_{i0}\right)X_t \tag{7}$$

$$\bar{x}_{ibt} = (\eta_{0b} + \epsilon_{ib})X^t \tag{8}$$

Where: η_{00} is a K-row vector that contains the overall coefficients that capture changes in attention over time, η_{g0} are deviations from η_{00} for consumers assigned to condition g (for identification, $\eta_{g0} = 0$ for $g = 5$), $\epsilon_{i0} \sim N(0, \Sigma_{10})$ is consumer unobserved heterogeneity; Σ_{10} is a diagonal matrix; X_t is a K vector that contains time scores corresponding to time t; $X_{tk} = (t-1)^{k-1}$ for $k = 1, \ldots, K$; η_{0b} are the effects specific to the b^{th} brand on display (for identification, $\eta_{0b} = 0$ for $b = 4$), $\epsilon_{ib} \sim N(0, \Sigma_{01})$ is brand specific unobserved heterogeneity; and Σ_{01} is a diagonal matrix.

3.4.2 Utility Accumulation

The model operationalizes the timing of attention to a brand ($\bar{\bar{x}}_{ibt}$) using two measures: (1) the time that has passed since the previous moment when the consumer fixated on the brand (d_{ibt}) and (2) a dummy variable that indicates if the brand is fixated on at time t ($I(fix_{ibt})$). Therefore, the utility of brand b at time t is reflected in the amount of attention ($\bar{x}_{ibt}$) the consumer has acquired thus far and the two measures that reflect its timing:

$$u_{ibt} = \alpha_b + \beta_1 \bar{x}_{ibt} + \beta_2 d_{ibt} + \beta_3 I(fix_{ibt}) + \varepsilon_{ibt} \tag{9}$$

Where: α_b are brand dummies that account for market level preferences for brand b, d_{ibt} is the time that has passed since the previous fixation on brand b, and $I(fix_{ibt})$ is an indicator function that takes the value 1 if the consumer inspects brand b at time t and zero otherwise; and ε_{ibt} is an unobserved utility shock.

Assuming that ε_{ibt} is type I extreme value distributed, the probability of choosing brand b as a function of the known part of brand utility U_{ibt} ($u_{ibt} = U_{ibt} + \varepsilon_{ibt}$) is:

$$p_{ibt} = \frac{\exp(U_{ibt})}{\sum_{b=1}^{B} \exp(U_{ijt})} \tag{10}$$

As discussed in section 1.3.1, there are different types of uncertainty that the consumer can experience during brand choice. In this chapter, we include an entropy measure that captures type IV uncertainty (about which of the brands in the set is best). In Chapter 5 we discuss alternative measures for this type of uncertainty and for the other three types. The extent to which consumer i is uncertain at time t is operationalized as:

$$H_{it}(\theta_t | x^t) = -\sum_{b=1}^{B} p_{ibt} \log p_{ibt} \tag{11}$$

Where H_{it} varies between 0 (when one brand has a choice probability of 1) and ln B (when the brands have equal choice probabilities).

We specify the utility consumer i derives at time t from continuing the search (u_{i0t}) as a function of an overall threshold (γ_0), uncertainty (H_{it}), and the time spent thus far (t):

$$u_{i0t} = \gamma_0 + \gamma_1 H_{it} + \gamma_2 t + \varepsilon_{i0t} \tag{12}$$

Where ε_{i0t} is an unobserved utility shock.

3.4.3 Moment-to-moment Utility Comparison

Assuming that the random utility term in equations 9 and 12 (ε_{ibt} and ε_{i0t}) is type I extreme value distributed, the probability of observing action $a_{it} = 0$, at time t is:

$$\Pr(a_{it} = 0|\alpha, \beta, \gamma) = \frac{\exp(U_{i0t})}{\exp(U_{i0t}) + \left(\sum_{b=1}^{B} \exp(U_{ibt})\right)^{\rho}} \tag{13}$$

Where ρ accounts for correlations between brand choice probabilities (Wooldridge 2010).

If the consumer decides to end the search, then the probability of choosing brand b^* is:

$$\Pr(a_{iT_i} = b^*|a_{iT_i} \neq 0, \alpha, \beta, \gamma) = \frac{\exp\left(u_{ib^*T_i}\right)}{\sum_{b=1}^{B} \exp\left(u_{ibT_i}\right)} \tag{14}$$

3.4.4 Model Estimation and Out-of-sample Predictive Performance

We use $D_i^{T_i} = (y_i^{T_i}, a_i^{T_i})$ to indicate the eye-fixations and choice data for participant i. The likelihood for participant i who expresses brand choice after T_i brand visits is:

$$\mathcal{L}_i(D_i^{T_i}) = \mathcal{L}^{fix}\left(y_i^{T_i}|\Theta\right)\mathcal{L}^{choice}\left(a_i^{T_i}|\Phi\right) = p\left(y_i^{T_i}|\Theta\right)\prod_{t=1}^{T_i} p(a_{it}|\Phi) \tag{15}$$

Where Θ is the collection of parameters that capture the link between eye movements and attention (eqs. 6-8) and Φ is the collection of parameters that describe the link between attention and choice (eqs. 9-14).

The model is estimated in R using RStan (R Core Team 2018; Stan Development Team 2018) using non-informative priors and multiple chains with dispersed starting values. Convergence is assessed using potential scale reduction (Gelman and Rubin 1992). The supplementary material (available on GitHub) includes all the code and data necessary to reproduce the results included in this chapter.

We compare model performance based on predictive accuracy calculated using K-fold cross-validation (Vehtari, Gelman, and Gabry 2017). The dataset is partitioned into $K = 11$ subsets D_k, for $k = 1, \ldots, K$. Then, the model is fit on a training dataset D_{-k} which contains all data for participants not part of D_k, and the posterior simulations are used to make predictions for participants in subset k. This approach mimics reality: a company has data about customers purchasing products which is used as input to determine market popularity and to calibrate the parameters of the model. Subsequently, this information is used to forecast what other customers will choose and when. Specifically, the model estimates Θ and Φ using observed data D_{-k}. Then, for participants in D_k it predicts the sequence of eye-fixations $\hat{y}_i^{T_i}$, the timing of choice $\hat{T}_i$, and the final brand choice $\hat{a}_{i\hat{T}_i}$.

The prediction performance for eye-fixations data ($\hat{y}_i^{T_i}$) is assesed using the density of the normal distribution. The log predictive density for the sequence of eye-fixations of participant i in subset k is:

$$\log p(\hat{y}_i^{T_i}|D_{(-k)}) = \ln \int p\left(\hat{y}_i^{T_i}\middle|\Theta\right)\mathrm{p}(\Theta|\mathrm{D}_{(-k)})d\Theta \tag{16}$$

Which is approximated by the expected log predictive density over S simulation draws from the posterior distribution $p(\hat{y}_i^{T_i}|D_{(-k)})$:

$$\widehat{elpd}_i = \log(\tfrac{1}{S}\textstyle\sum_{s=1}^{S} p(\hat{y}_i^{T_i}|\Theta^{\mathrm{k,s}})) \tag{17}$$

The expected log predictive density for model m over all participants $i = 1, \dots, N$ is:

$$\widehat{elpd}_m = \textstyle\sum_{i=1}^{N} \widehat{elpd}_i \tag{18}$$

We compare the predictive performance of two models by calculating the difference in expected log predicted density ($\widehat{elpd}_{m1} - \widehat{elpd}_{m2}$) and standard error of this difference:

$$se\left(\widehat{elpd}_{m1} - \widehat{elpd}_{m2}\right) = \sqrt{N * Var_{i=1}^{N}(\widehat{elpd}_{m1i} - \widehat{elpd}_{m2i})} \tag{19}$$

The model predicts the timing of choice $\hat{T}_i$ and the final brand choice $\hat{a}_{i\hat{T}_i}$ for each participant from moment-to-moment, as more eye movements are observed. Hence, for participant i the model makes T_i predictions. The predictions made at moment t use the eye movements observed for that participant so far (y_i^t) and the estimated parameters based on participants in the training sample D_{-k}. First, we extract $\hat{\eta}_{it}$ from the sequence of observed eye movements y_i^t. Then, we use this to simulate search and brand utilities at future moments: $\hat{u}_{i0v}$ and $\hat{u}_{ibv}$ where $t \leq v$ and calculate the probability that the participant continues inspecting the brand for all moments v where $t \leq v$. The predicted timing of choice $\hat{T}_i$ is calculated by integrating over these probabilities. While in theory the integral should be taken over an infinite time horizon it is unlikely that participants continue to inspect the brands indefinitely. We found that using a time horizon two times longer than the maximum decision time observed in the study ensures the probability of continuing the process converges to zero. Hence, the moment of choice is predicted as:

$$\hat{T}_i = t + \int_t^{100} (1 - \hat{p}_{i0v}) dv \tag{20}$$

Given the predicted moment of choice $\hat{T}_i$ and the brand utilities at that moment, the model predicts the brand that the participant is going to choose based on the corresponding brand choice probabilities. The predictive performance for brand choice ($\hat{a}_{i\hat{T}_i}$) predictions (PP_m^{BC}) for model m is based on the brand choice hit rate. This is calculated as the average predicted choice probability of the chosen brand (b^*) at the predicted moment of choice ($\hat{T}_i$):

$$PP_m^{BC} = \frac{1}{N} \sum_{i=1}^{N} \hat{p}_{ib^*\hat{T}_i} \tag{21}$$

The predictive performance for moment of choice ($\hat{T}_i$) predictions (PP_m^{MC}) for model m is calculated as one minus mean absolute percent error:

$$PP_m^{MC} = 1 - \frac{1}{N} \sum_{i=1}^{N} \frac{|T_i - \hat{T}_i|}{T_i} \tag{22}$$

3.5 Results

Model free evidence shows that participants allocate a larger share of fixations on the brand that they eventually choose compared to participants who choose a different brand (Figure 3.5). For example, brand 1 has a 49% fixation share for participants who choose it but only 32% for participants who choose a different brand. The difference in fixation shares per brand between participants who choose the brand and those who do not increases from moment-to-moment until choice is expressed. This suggests that there is heterogeneity in attention at the participant-level and at the brand-level. Similar to chapter 2, there is empirical support in the eye-movement data for heterogeneity in attention both between participants and brands. Appendix C presents the results of comparisons between the model specified in equations 6-8 and alternative models that restrict brand or participant-specific attention to zero. These results (Appendix C) support the idea that consumers strategically inspect the brands (model component 1 in Table 3.1). Table 3.3 presents estimates for the amount of heterogeneity at the participant and brand levels. The results show more heterogeneity at the brand-level as compared to participant-level heterogeneity. There are differences both in the initial attention to the brands (eq. 8, σ_{011} = .46, CI = [.44; .49]) and in how brands are inspected from moment-to-moment (eq. 8, σ_{012} = .74, CI = [.69; .79] and σ_{013} = .32, CI = [.29; .36]). If participants were to inspect all the brands in a similar way, for example by processing the information by brand or by attribute, then it would not be possible to infer any differences in utilities before observing choice.

3.5.1 Model Comparison

As described in section 3.2.6, the proposed model (M6) is compared against two groups of competing models. We compare these models in terms of predictive performance first at the moment of choice, when all the eye movements are observed, (Table 3.2) and in section 3.5.4 present moment-to-moment predictive performance. We discuss implications for the link between attention and utility accumulation, and for determinants of choice timing.

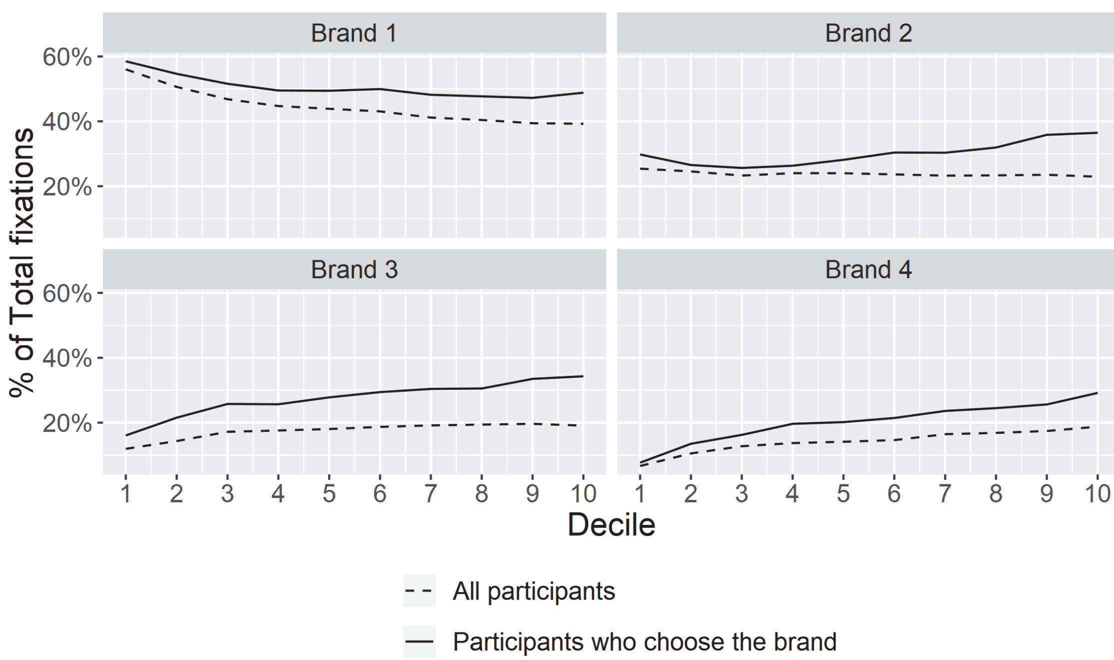
First, we compare the proposed model against M1 and M2. Because these models use different specifications for the determinants of *what* brand to choose, we expect to see differences in predictive performance both for brand choice and moment of choice. Even if M1, M2, and the proposed model use the same determinants of *when* to express brand choice,

Table 3.3 Estimated heterogeneity in information acquisition

Parameter		M	SD	p-value	2.5%	97.5%
Overall growth	η_{001}	-.05	.04	.21	-.13	.03
	η_{002}	2.67	.10	<.001	2.46	2.87
	η_{003}	-1.24	.10	<.001	-1.42	-1.05
	η_{004}	.34	.04	<.001	.26	.42
	η_{005}	-.03	.01	<.001	-.04	-.02
Participant Heterogeneity	σ_{101}	.03	.02	<.001	.01	.07
	σ_{102}	.05	.02	<.001	.01	.10
	σ_{103}	.03	.01	<.001	.01	.06
	σ_{104}	.01	.004	<.001	.001	.02
	σ_{105}	.002	.001	<.001	.001	.005
Brand Heterogeneity	σ_{011}	.46	.01	<.001	.44	.49
	σ_{012}	.74	.03	<.001	.69	.79
	σ_{013}	.32	.02	<.001	.29	.36
	σ_{014}	.02	.01	<.001	.01	.04
	σ_{015}	.01	.001	<.001	.01	.01
Unobserved Heterogeneity	σ_{00}	.28	.002	<.001	.27	.28

Note: M = Mean; *SD* = Standard deviation; 2.5% and 97.5% = percentiles of the posterior distribution; *p*-value = Bayesian one-tailed p-value. Participant and brand specific effects are not shown to save space.

Figure 3.5 Participants fixate more on the brand they eventually choose



the search utility is compared against different brand utility values (due to different specifications of brand utility). M1 specifies that the mapping from brand utilities to eye-fixations occurs without error ($\xi_{it} = 0$, eq. 1), thus using the observed number of eye-fixations to specify brand utility. While M2 accounts for imperfect mapping between what utility and eye-fixations (ξ_{it}, eq. 1), it assumes that the moments when the consumer inspects the brands do not reflect brand utilities over and above the amount of attention. This competing model restricts β_2 and β_3 in equation 9 to zero. These differences in the specification of brand utility lead to different predictions for the moment of choice.

M1 and M2 differ in how they account for the link between eye movements and attention. By comparing these models, we find that it is important to account for unobserved heterogeneity in eye movements and extract the brand and participant specific attention measures ($\bar{x}_{ibt}$) as opposed to using the observed number of fixations. While the brand choice hit rate at the moment of choice is better for M1 (57.2% vs 54.5%) the prediction accuracy for the moment of choice is much lower (53.5% for M1 vs 72.7% for M2). The difference in predictive performance between M1 and M2 suggests it is important to account for error in the mapping between utility and eye movements (ξ in eq. 1), especially if we want to understand what determines the moment when choice is expressed.

Comparing the brand choice hit rate at the moment of choice, we see that the proposed model has a hit rate of 59.6% which is larger than that of M1 (57.2%) and M2 (54.5%). Different from the proposed model, M2 assumes that brand utilities are not influenced by the timing of attention. The difference in brand choice hit rate at T_i between M2 (54.5%) and the proposed model (59.6%) shows that it is important to account for when the consumer focuses on a brand, in addition to the amount of attention on the brand. This is consistent with previous studies that find increasing number of fixations on the chosen brand in the final few moments before choice is expressed (Atalay et al. 2012; Shimojo et al. 2003).

The second group of models (M3-M5) makes different assumptions about the determinants of *when* to express brand choice. These models all use the same specification for brand utility, so the minor differences ($\pm$ .1 pp) in brand choice hit rate at the moment of choice are due to sampling variation in the MCMC chain. Therefore, we focus on interpreting results for the accuracy of moment of choice predictions. M3 (62.3%) has a lower performance as compared to M4 (67.6%), M5 (73.3%), and the proposed model (73.4%) at the moment of choice. Similar to M3, M4 also assumes that consumers compare the search utility against a fixed decision threshold, but allows consumer uncertainty to influence search utility and the moment when brand choice is expressed. By comparing M4 against M3 we find that accounting for consumer uncertainty improves moment of choice predictions by about 5pp. However, the improvement in moment of choice predicts disappears for models that specify that consumers compare search and brand utilities. Specifically, the predictive performance of the proposed model (73.4%) is very similar to that of M5 (73.3%), which restricts the contribution of consumer uncertainty to zero.

By comparing the proposed model against M3 and M4, the results show strong support for participants comparing search and brand utilities from moment-to-moment, as opposed to comparing search utility against a fixed threshold from moment-to-moment and brand utilities against each other only at the moment of choice. If the model were to outperform M5, then this would have been evidence that consumers take into account not only the maximum brand utility they receive by ending the task, but also how uncertain they are about which brand is best. However, the results do not support this.

These results demonstrate that to understand *how* decisions are made, one needs to account for (1) *how much* consumers have inspected each brand, (2) *when* they have done so, and (3) the moment-to-moment comparison between search and brand utilities.

3.5.2 Utility Accumulation

Determinants of What Brand to Choose. Both the amount of attention and the moments when brands are inspected impact brand utilities. Specifically, brands that receive more attention are more likely to be chosen (2.02, p-value <.001). This supports the assumption that consumers allocate more attention to promising brands (Table 3.1: Eye movements). The time since the last visit on the brand has a negative effect on brand utility -.30 (p-value <.001). This implies that brands that have recently been fixated on have a larger choice probability. The results do not support a positive effect of current visit on brand utility (.41, p-value = .06). The estimated brand intercepts are in line with observed brand choice shares and fixations frequencies, as brands 2 and 3 are chosen by fewer participants than brands 1 and 4. While brands 1 and 4 have similar choice shares, brand 1 receives more fixations. The positive intercept for brand 4 (1.60, p-value <.001) reflects this. Overall, the results are consistent with Chapter 2 and show that those brands that the participant inspects in more detail accumulate more evidence, especially if this takes place over repeated visits closer to the moment of choice.

Determinants of When to Express Brand Choice. The overall threshold (γ_0 = 8.06, p-value < .001) is the level of brand utility required to stop the process after the first visit. The decision threshold decreases due to costs of time (γ_2 = -.15, p-value < .001). On average, search utility decreases by one unit of utility every seven brand visits. The estimated cost of time corresponds to a reduction of 3.15 units of utility in the decision threshold for the average duration of a decision in our sample (21 brand visits). If the decision to express brand choice were influenced by consumer uncertainty, then γ_1 would be positive. However, the results do not support this (-.09, p-value = .984).

Table 3.4 Attention reflects why consumers choose *what when*

		M	*SD*	*p*-value	2.5%	97.5%
Determinants of *What* brand to Choose						
Brand intercepts						
Brand 2	α_2	-.25	.28	.33	-.77	.30
Brand 3	α_3	-.70	.35	.04	-1.41	-.06
Brand 4	α_4	1.60	.29	<.001	1.05	2.16
Attention:						
Amount	β_1	2.02	.27	<.001	1.49	2.54
Timing - Time since last visit on brand	β_2	-.30	.09	<.001	-.49	-.13
Timing - Current visit on brand	β_3	.41	.26	.06	-.10	.91
Determinants of *When* to Express brand Choice						
Overall threshold	γ_0	8.06	.73	<.001	6.57	9.46
Uncertainty	γ_1	-.09	.37	.984	-.85	.63
Cost of time	γ_2	-.15	.02	<.001	-.18	-.11

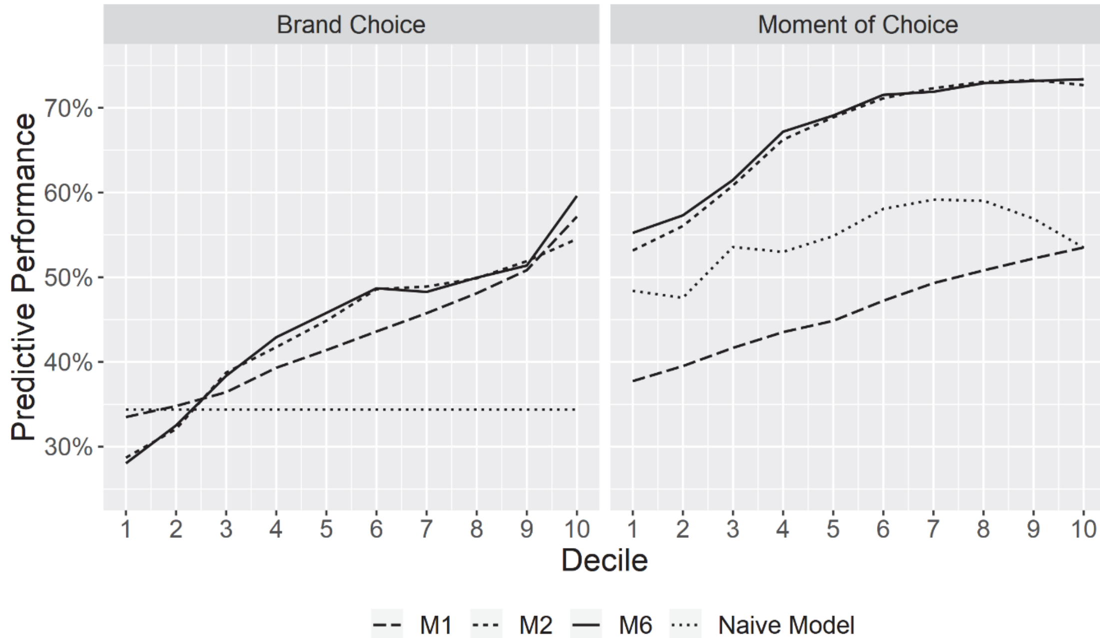
Note. M = Mean; *SD* = Standard deviation; 2.5% and 97.5% = percentiles of the posterior distribution; *p*-value = Bayesian one-tailed *p*-value. Brand 1 is the reference brand.

3.5.3 Sequential Predictions for Brand Choice and Moment of Choice

The model predicts what brand consumers are going to choose and the moment when they express their choice. Both predictions are out-of-sample and from moment-to-moment. As more eye-fixations are added to the x_i^t sequence, the model updates the attention measures (amount and timing of attention). These are used to simulate brand and search utilities at future visits ($t < v$) and predict the consumer's actions, as described in section 3.4.4. We present the accuracy of these predictions summarized by decision decile.

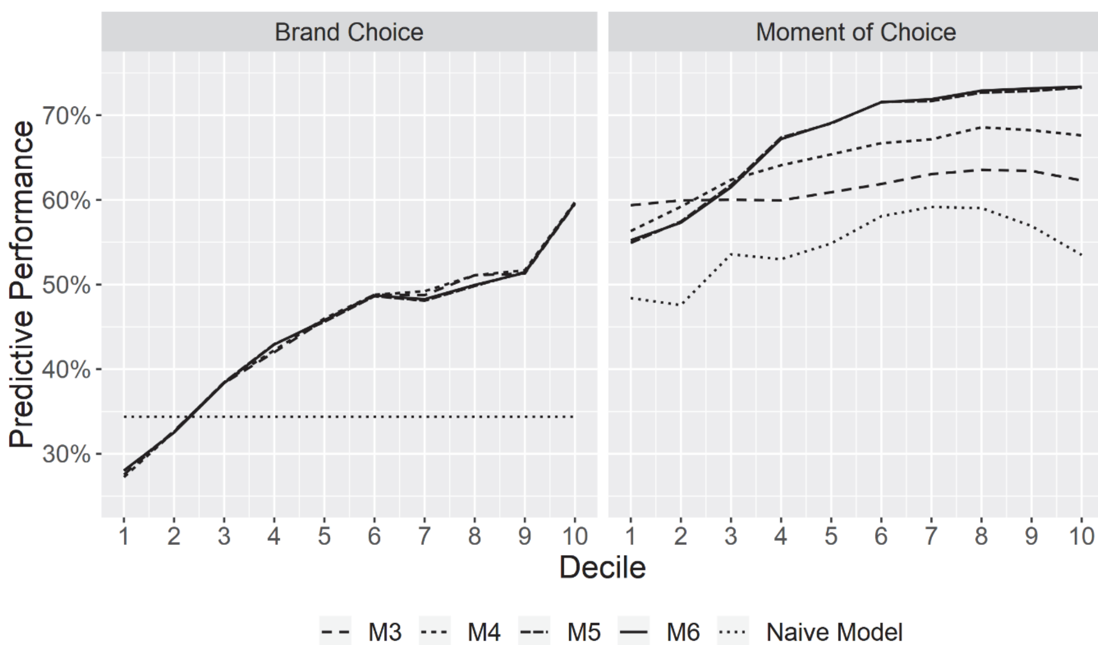
The brand choice hit rate reaches 59.6% at the end of the process. A naïve model that predicts brand choice based on observed brand choice shares in the calibration sample has a hit rate of 34.39%. Starting with the third decile our model is already above this hit rate baseline. In the beginning of the decision process the model underestimates how much time consumers need and predicts that they stop earlier than they actually do. However, brand choice and choice timing predictions become more accurate over time (Figures 3.6 and 3.7). This evidence supports the assumptions we make about the components 2, 3, and 4 of the model (Table 3.1), namely that moment to moment search and brand utilities predict what consumers are going to choose and when. And that these predictions improve over time.

Figure 3.6 Comparison of out-of-sample prediction accuracy for models with different determinants of *what* brand to choose



Note: Naïve model predictions for brand choice are based on observed brand choice shares in the calibration sample. Naïve model predictions for moment of choice are based on the distribution of decision times in the calibration sample.

Figure 3.7 Comparison of out-of-sample prediction accuracy for models with different determinants of *when* to express brand choice



Note: Naïve model predictions for brand choice are based on observed brand choice shares in the calibration sample. Naïve model predictions for moment of choice are based on the distribution of decision times in the calibration sample.

3.6 Discussion

The findings of this chapter provide insights into fundamental decision processes that describe how consumers determine (1) *what* brand to choose and (2) *when* to express their choice, and (3) how these two decisions are closely linked to the sequential eye movements that consumer use to inspect the brands. While it builds on previous results for the overall link between brand choice and eye movements (Glaholt et al. 2009; Krajbich et al. 2010; Pieters and Warlop 1999), to the best of our knowledge, the proposed model is the first to jointly investigate the drivers of both brand choice and choice timing when consumers make a single preference-based choice between multiple complex alternatives.

Table 3.1 summarizes the components of the model developed in this chapter and the assumptions that it is based on. For each of these assumptions we formulate implications and indicate the empirical evidence that is needed to support them. To support all the characteristics and assumptions of our model we would need 11 pieces of evidence, which consist of (1) estimated values for parameters of interest and (2) measures of out-of-sample prediction performance for brand choice and moment of choice. We find empirical support for nine of the 11 required pieces of evidence. The results indicate that the timing, amount, and sequence in which consumers focus on the brands reflect utility accumulation throughout the decision process. This supports the main proposition of our model – that differences in attention between the brands build up over time as the consumer evaluates their utilities based on sequentially acquired information. We now summarize the main results and in the next section discuss their implications.

As in chapter 2, we find (1) empirical support for heterogeneity in attention and (2) that consumers attend more to the brand they are more likely to choose. Even though all participants were presented with the same information, they inspected the brands in different ways. The one similarity between how participants examine the brands is that they focus more on the brand they eventually choose. Model free evidence and estimation results show that this happens regardless of the brand consumers eventually choose. This supports the idea that utilities build up over time and are updated in light of newly acquired information about the brands. These results are important because brand-heterogeneity in attention makes it possible to infer brand utilities from moment to moment in the absence of any other prior preference measurements.

Importantly, and novel, the proposed model predicts *when* consumers are going to express their choice, in addition to *what* brand they choose. This is evidence of a fundamental link between moment-to-moment eye movements and unobservable processes of attention and utility accumulation that result in brand choice.

Second, accounting for unobserved sources of heterogeneity in eye movements improves predictions, in particular moment of choice predictions. This implies that observed fixations do not accurately reflect brand utility, despite making reasonable predictions of what brand consumers are going to choose. More specifically, the accuracy of brand choice predictions based on the observed number of fixations are surprisingly similar to those of the proposed model. This means that models that use the number of fixations, as it is common in the literature, can still provide insights into which brand consumers like best. However, they make poor predictions (M1) of the moment when consumers express their choice.

Third, the results suggest that consumers compare brand utilities against a time-varying decision threshold. Hence, it is important for brands to attract consumers' attention early on as this can terminate search faster before competing brands steal the spotlight and sway consumers' choice.

Fourth, both the moment of choice and brand choice can be predicted early on and more importantly, without having any information about the participant (e.g. prior choices, preference ratings, brand ownership). Even more, the accuracy of these predictions improves over time as the consumer inspects the brands.

3.6.1 Implications and Future Work

This research extends prior work that has primarily focused only on *what* brand consumers choose, but that has not explored determinants of *when* consumers express brand choice. This applies both to empirical applications (Atalay et al. 2012; Glaholt et al. 2009; Meißner et al. 2016; Pieters and Warlop 1999) and analytical developments of RIT (Matějka and McKay 2015; Steiner et al. 2017) which so far have not been tested empirically. While previous SSMs such as the aDDM (Krajbich et al. 2010) use decision duration to explain the computational processes that take place during choice, it is not immediately obvious if and how they can be adapted to typical consumer choices. More specifically, applications of the aDDM use multiple repeated choices between at most three single-attribute brands while consumers very often make single choice between multiple multi-attribute brands for which no prior preference measurements are available.

Developments in the RIT literature thus far focus primarily on analytical models that formalize how decision makers choose what information to acquire and the impact of this choice on the outcome of the decision process (Matějka and McKay 2015; Steiner et al. 2017). Such models are based on the assumptions that consumers optimally choose what information to acquire. However, it is unclear to what extent optimal information acquisition, which implies the exclusive use of top-down directed attention, can take place in realistic decision settings. For example, online retailers use advertising, product recommendations, and changes in webpage layout to attract consumers' attention and to influence what information they are exposed to, with the final goal of influencing their purchase behavior. Therefore, it is probably more realistic to accommodate for a mix of bottom-up and top-down directed attention. Our model offers this possibility by extracting brand- and participant-specific attention measures. The proposed model further relaxes the RIT assumption of optimal information acquisition by also accounting for unobserved random error in the mapping from utility to eye movements (ξ eq. 1).

The results of this study suggest three directions for future research. First, it would be interesting to investigate if eye movements data can be used to improve preference measurement in situations when repeated choices cannot be observed or when there is reason to believe that observed choices are not an accurate reflection of consumers' true preferences. For example, in situations when consumers have an incentive to be dishonest or feel pressured to give a certain answer (e.g. attitudes towards substance use, unhealthy lifestyle choices, political preferences). Standard choice models assume that consumer preferences are revealed by their final stated choice and require multiple observed choices in order to estimate consumer heterogeneity in preferences. Therefore, they are not well suited in these situations. The model developed in this chapter offers a starting point because it uses brand and consumer specific attention to infer utilities before the final choice is observed.

Second, the introduction indicates that the proposed model can be adapted to use other types of data that capture repeated consumer interactions or touchpoints. This is useful for decisions that are taken over a longer time and in multiple sessions, when eye-tracking participants is no longer an option. Examples of other types of consumer level data that can be used with the proposed model are: browsing history and geolocation data. Both types of data are easily accessible to companies who use cookies to collect consumer browsing behavior or mobile apps that track GPS coordinates in real time. We argue this would make the application of our model relevant also for physical stores, not only for online retailers.

Third, Chapter 2 discussed how the proposed model can be trained on eye-movement and choice data of participants and then be used to predict choice for new participants. The model in this chapter implements and tests this idea. This is one step further towards influencing consumer choice. The next step is to look into how and when it is best to intervene and try to influence *what* consumers choose or *when* they express choice, or both. For example, if the participant appears to favor a brand, then an intervention could influence him or her to express choice faster to prevent the discovery of competing brands. But if the consumer favors a competitor's brand than an intervention could facilitate the discovery of other brands with the goal of changing the final brand choice. Such interventions, while obviously beneficial for companies, raise multiple ethical concerns. These ethical concerns are even more important if the choices are in the health domain, involve under-age decision makers, or have consequences for political elections. Therefore, as methods and models are developed to influence choice, it is necessary to also focus on limiting or restricting their use. We come back to this point in chapter 5.

Several topics for discussion remain. Though our model controls for brand effects on attention, the better option would have been to just randomize the order of the brands. The goal manipulations used in our study did not impact attention or choice which limits what our results can say about the role of decision goals on attention and choice, or both. Both limitations of our study can be addressed by future research that uses incentive aligned studies and randomizes the order in which brands are presented on the display.

Second, our results focus on choices from a preset choice set. While this is a frequent choice situation, the choice set is commonly also endogenously determined. Studies that account also for the decision to include a brand in the final choice set might find that this also reflect consumer preferences and are expected to have an even better brand choice hit rate.

Current models of information acquisition and choice are well suited to measure preferences post-choice. However, to understand *how* preferences are constructed within a single choice, a new approach is needed. This chapter offers such an approach and thus provides insights into how consumers arrive at their choice and the opportunity to make moment-to-moment predictions for *when* consumers choose *what*.

Chapter 4

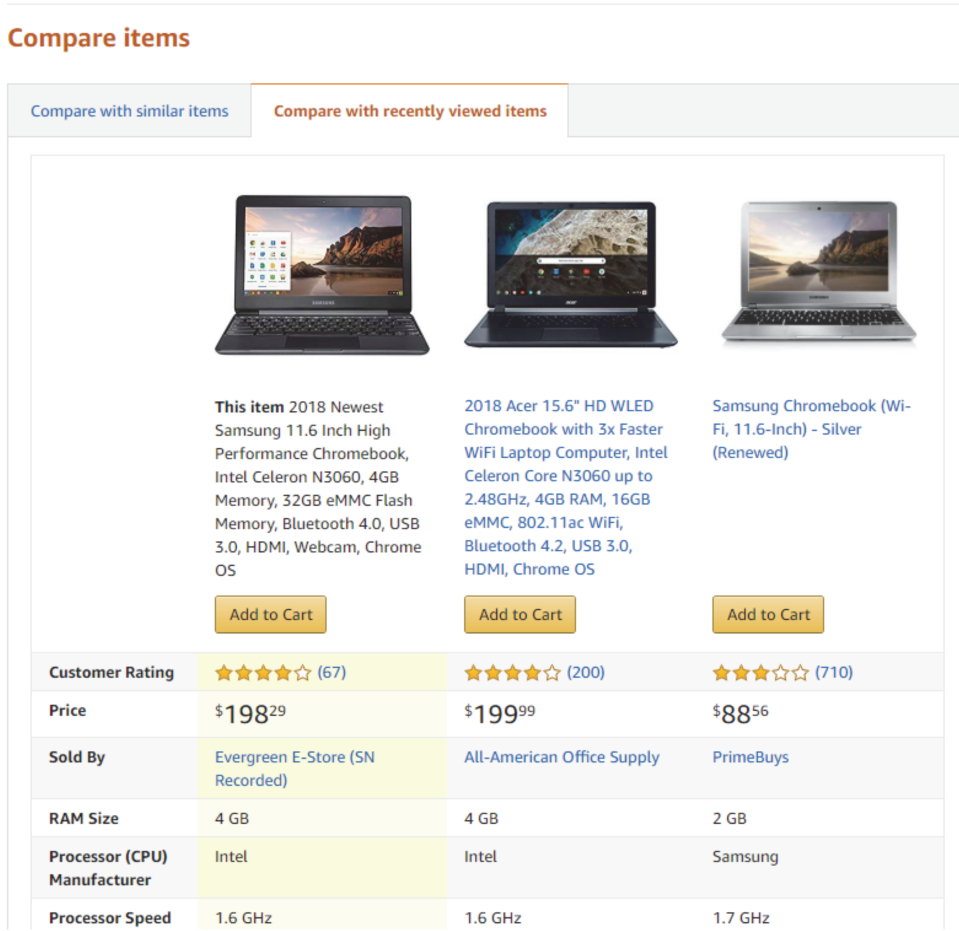
Attention, Attribute Importance, and Brand Choice

4.1 Introduction

This chapter investigates the idea that eye movements reflect both what brand consumers are going to choose, and why this brand is preferred over other alternatives in the set. We develop a model that infers attribute importance from how consumers allocate their attention over time to the information on display before expressing brand choice. This builds on theoretical perspectives of rational inattention theory (RIT) (Caplin and Dean 2015; Steiner et al. 2017) and multialternative decision field theory (MDFT) (Roe et al. 2001), and generalizes the models developed in chapters 2 and 3. While analytical RIT models focus on the overall link between attention to a brand and its utility, the model developed in this chapter adapts RIT theoretical assumptions to choice between brands described by multiple attributes. The models proposed in Chapters 2 and 3 extract attention measures that are linked to overall brand utilities. In this chapter, we take a different approach and decompose brand utilities into two components that capture the importance of the attributes that describe the brands and the subjective value that the consumer attaches to the attribute levels corresponding to each of the brands on display. This implies that eye movements reflect not only how the consumer evaluates the brands, but also why some brands are preferred.

Consumers frequently make online purchases for one brand from a set of competing alternatives. Shopping websites facilitate search for a brand that matches their preferences among thousands of competing alternatives, each of them described by multiple features (i.e. attributes). For example, Amazon users can use filters to view only brands that meet certain criteria (e.g. price range) and are presented with information about brands in a format that facilitates comparison (Figure 4.1). This allows them to narrow down the search to a set of brands that they inspect by moving their eyes to areas of the display corresponding to specific attributes.

Figure 4.1 "Compare with recently viewed items" section at amazon.com



Note: retrieved June 2019[6].

While consumers could inspect all the brands and attributes on the screen, it is in their interest to focus on information that is most relevant for their choice (Caplin and Dean 2015). This has three implications. First, preferred brands are examined more, as already supported by the results of chapters 2 and 3, and previous research that examines eye movements during

[6] https://www.amazon.com/Samsung-Performance-Chromebook-Celeron-Bluetooth/dp/B0747QTRN4/ref=sr_1_6?keywords=samsung+chromebook&qid=1561905545&s=gateway&sr=8-6

brand choice (Chandon et al. 2009; Glaholt et al. 2009; Pieters and Warlop 1999). Second, consumers allocate more attention to important attributes in order to identify the brand that performs best on specific dimensions. For example, a price sensitive consumer is expected to focus more on price and less on other brand attributes as compared to what would be expected if all attributes were equally important. Third, the moments when attributes are attended to provide additional information about how important they are for the consumer. Because inspecting the brands takes time and effort, consumers are expected to focus on the most important attributes already early in the process (Russo and Rosen 1975; Wedell and Senter 1997). Then, they can examine additional attributes that are not as important and closer to the moment of choice, consumers can go back and briefly check important information as part of a verification stage (Russo and Leclerc 1994). These three implications suggest that attribute importance weights vary over time and are closely related to the moment-to-moment eye movements that consumers make during a single brand choice. This offers the possibility to extract time-specific attribute importance weights from moment-to-moment eye movements during a task.

Traditional methods for assessing attribute importance include: (1) asking consumers for direct ratings of importance, (2) using conjoint measurement techniques, or (3) open-ended elicitation (Jaccard, Brinberg, and Ackerman 1986). Even though these methods aim to measure the same construct, namely attribute importance, they can produce contrasting results (Heeler, Okechuku, and Reid 1979). Choice-based conjoint (CBC) studies use repeated choices in the same product category to estimate consumer specific subjective values (partworths) of different attribute levels (Green and Srinivasan 1990). Then, the importance of each attribute is calculated as a ratio between the range of estimated partworths for that attribute relative to the sum of the partworth ranges for all attributes (Green and Srinivasan 1978). The number of choices that need to be observed before measuring attribute importance increases as a function of the number of attributes and attribute levels that describe the brands. To the extent that consumers do not examine all the information on display, augmenting choice with eye-fixation data lowers the necessary number of repeated choices per participant in conjoint studies (e.g. from 16 to 12 choices per participant (Yang et al. 2015)). Nevertheless, multiple choices within the same category are required in conjoint studies and the attribute importance measures that are derived are agnostic to the moments when the consumer has inspected them. However, consumers are more likely to inspect important attributes at earlier moments during choice (Payne, Bettman, and Luce 1996; Simonson, Huber, and Payne 1988). Therefore, the importance of an attribute is reflected by the moments when the attribute is attended to, in addition to the amount of attention it receives. Asking consumers for direct ratings of importance and open-ended elicitation methods also suffer from similar limitations (Bottomley, Doyle, and Green 2000; Zhu and Anderson 1991). The model developed in this chapter addresses these limitations and contributes in three important ways.

First, the model extracts consumer-specific attribute importance weights based on attention allocation during a single brand choice. Building on previous research (Bettman et al. 1998; Payne 1976; Pieters and Wedel 2007), the proposed model assumes that attribute importance weights are influenced by specific decision goals that consumers have. This implies that differences in attention between attributes are primarily motivated by top-down attention processes. The results of this chapter extend (1) prior work that has used information display boards to examine the relationship between information acquisition and importance weights (Wedell and Senter 1997), and (2) recent developments in economics that examine the role of importance weights in choice, but assume these weights are extrinsically determined by how much the attribute level stands out, i.e. is salient, as compared to the other brands on display (Bordalo et al. 2013; Bordalo, Gennaioli, and Shleifer 2016). If importance weights were choice set specific, as these models in economics assume (Bordalo et al. 2013, 2016), as opposed to consumer-and-time specific as in the model we propose, then two consumers making a choice from the same set would use the same importance weights. The results included in this chapter provide evidence against such simplifying assumptions and support our theory. Namely, that attribute importance weights are (1) consumer specific, (2) influenced by the decision goal of the consumer, and (3) vary over time.

Second, the model decomposes brand utility in two components: attribute importance and consumer's subjective values for the features of the brand. We do this for single brand choices and thereby extend previous theories and methods that require consumers to make repeated choices within the same category in order to estimate consumer-specific preferences. The results reveal that decision goals influence not only what attributes are attended, but also conditional on attending to an attribute, what brands are inspected and for how long. This extends MDFT models in two ways. First, MDFT models assume attention influences only the selection of which attribute is attended to at a specific time, but not how the brands are evaluated within that attribute (Roe et al. 2001). Second, empirical applications of MDFT make two simplifying assumptions (Diederich 2003; Dror, Basola, and Busemeyer 1999): (1) participants derive the same values from the different attribute levels, and (2) these values are equal to the corresponding attribute levels, which requires all attribute levels to be numeric. The model developed in this chapter infers participant specific values for different attribute levels from sequences of eye movements and can be calibrated for any type of attribute information (e.g. textual, numeric, pictorial).

Third, attribute importance and subjective values of brand-and-attribute combinations are extracted from moment-to-moment eye movements, and the model can accommodate any number of brands and attributes. This offers the possibility to make brand choice predictions based only on measures extracted from eye movements. The accuracy of these predictions is above chance (25%, one in four brands) and even above standard logit models. Specifically, the proposed model has an average brand choice hit rate of 63%, 26 percentage points larger than that of a logit model that accounts for observed brand choice shares (37%), and 8 percentage points larger than that of a logit model that includes information about the type of brand that the consumers aims to choose (55%). In addition, the results of this chapter establish the extent to which attention-based attribute importance weights differ from importance weights typically estimated in a choice model (M5 versus M3, sections 4.2.3 and 4.4.4). The results show that brand choice can be predicted even for new consumers in new categories. Of course, these predictions are also possible for existing customers, and we explain in section 4.2.2 how the model can include available information about consumers, brands, and attributes.

We calibrate our model on eye-tracking data collected in a study (*N* = 334) on single choice between four multi-attribute brands. The next section offers a brief description of the eye-tracking and choice study and introduces the proposed model. Then, we describe in detail the study design and eye-movement measures (section 4.3), and the results of the proposed model (section 4.4). The chapter ends with a discussion of the implications and limitations of the proposed approach.

4.2 Model and Related Literature

Before describing the model and the related literature, this section introduces the eye-tracking and choice study. Participants were instructed to (1) choose the most environmentally friendly brand (eco-goal, $X^{goal} = 1$) or (2) choose the brand that performs the best (performance-goal, $X^{goal} = 0$) from a set of four brands. Decision time was externally controlled: (1) 15 seconds (high time-pressure, $X^{tp} = 1$) or (2) 30 seconds (low time-pressure, $X^{tp} = 0$). The study used a between-subjects factorial design and random assignment to conditions. After being reminded of the decision goal and available time to inspect the brands, participants were presented with an attribute-by-brands matrix that contained information on several attributes for four brands. Participants made single choices in four product categories, while their eye movements were recorded. For each participant (i) and choice task (c), the decision time is split in three intervals (t). This facilitates comparisons between participants with in different time-pressure conditions, as each time interval corresponds to one third (5s for participants in the low time-pressure conditions and 10s for participants in the high time-pressure condition) of their respective decision time. One observation corresponds to the number of eye-fixations on brand b and attribute j for participant i, choice task c, and interval t. Two measures are derived: (1) per attribute, the share of eye-fixations as percent of the total number of fixations (y_{ci0jt}) and (2) per brand and attribute, the share of eye-fixations on the brand as percent of the number of fixations on the attribute (y_{cibjt}).

In order to draw general inferences about the effects of decision goals and time-pressure on attention towards different types of information on display, attributes are grouped in three categories (Roe et al. 2001). Specifically, the three ($J = 3$) types of attributes are: (1) eco-attributes ($j = 1$) that provide information needed to select the most environmentally-friendly brand, (2) performance-attributes ($j = 2$) that provide information needed to select the brand that performs the best, and (3) other-attributes ($j = 3$) that provide additional information not directly relevant to either of the two decision goals. For example, in the third choice task between four TV brands, eco-attributes are energy class, on-mode power consumption, and electricity consumption for a year; performance attributes are image motion rate, image quality, resolution, and audio power; and other-attributes are brand name, screen size, and dimensions. The names of all attributes and their classification into attribute types for each of the four product categories is included in Table 4.1.

4.2.1 Brand Utility

This chapter develops a model to describe consumer choice from a set of B brands for which information about J attributes is displayed at the same time, in an attribute-by-brands matrix. Consumers use their eyes to inspect the information on display and are strategic in their choice of what brands and attributes to inspect from moment-to-moment. Building on RIT (Caplin and Dean 2015) and multialternative decision field theory (MDFT) (Roe et al. 2001), consumers allocate more attention to promising brands and important attributes.

The utility that a participant i derives from choosing brand b described by J attributes accumulates over time as the participant inspects the information on display. Specifically, brand utility is influenced by (1) the importance of the attributes (β_{ijt}) and (2) the subjective value (θ_{ibjt}) that consumer i assigns to the features of brand b. Consider a participant aiming to choose the most environmentally friendly TVs from a set of four brands. The participant forms a subjective value for each of these brands that reflects the extent to which their features match the eco-goal of the consumer. For example, TV brands with lower energy consumption offer more value to this specific participant. Because a larger weight is placed on eco-attributes such as energy class and consumption (captured by β_{i1t} ($j = 1$ (eco)), the participant is better able to discriminate between the brands on this particular attribute, even for small differences between their respective subjective values (θ_{ibjt}). As the participant examines the information on display, the utility of the brands changes. The utility of brand b at time t for participant i is:

$$u_{ibt} = \sum_{j=1}^{J} \beta_{ijt} * \theta_{ibjt} + \varepsilon_{ibt}, \tag{1}$$

where β_{ijt} is the importance of attribute j for participant i at time t; θ_{ibjt} is the subjective value of brand b on attribute j for participant i at time t; and ε_{ibt} is an unobserved random utility shock. All model parameters are choice task specific but to simplify notation, subscript c is omitted.

Table 4.1 Choice task description: product category, brands that match participant decision-goals, and attribute types

Choice task	Product Category	Eco-friendly Brand	Best-performance Brand	Number of attributes J_c	Eco-attributes (f_{cj})	Performance-attributes (f_{cj})	Other attributes (f_{cj})
1	Light bulb	B, D	C	9	bulb type, energy class, wattage, lifetime (44%)	light output, watt equivalent (22%)	brand name, voltage, color (33%)
2	Travel Mug	D	A, C	6	recycled (17%)	material (17%)	brand name, volume, weight, size (66%)
3	TV	C	B	10	energy class, on-mode consumption, consumption per year (30%)	image quality, image motion rate, resolution, audio power (40%)	brand name, screen size, dimensions (30%)
4	Fridge	B	A, D	10	energy class, consumption per year (20%)	fast chill option (10%)	brand name, cooling space, freezing space, compartments, freezer position, size, weight (70%)

Assuming that ε_{ibt} is extreme value type I distributed, and $U_{ibt} = \sum_{j=1}^{J} \beta_{ijt} * \theta_{ibjt}$, the probability that participant i chooses brand b at time t is:

$$p_{ibt} = \frac{\exp(U_{ibt})}{\sum_{b=1}^{B} \exp(U_{ijt})}. \tag{2}$$

Because consumers allocate more resources to information that is more important for their decision task or goal (MacKenzie 1986; Pieters and Wedel 2007; Wedell and Senter 1997), important attributes receive a larger share of attention ($\bar{\beta}_{ijt}$) as compared to what would be considered a fair share. Assuming similar levels on information complexity between the attributes on the screen, if all attributes were equally important to the participant then the fair-share is given by the number of attributes of a specific type divided by the total number of attributes (f_j in Table 4.1). The ratio between the attention share and this fair share captures the relative advantage or disadvantage of an attribute type ($\beta_{ijt} = \bar{\beta}_{ijt}/f_j$). The share of attention on an attribute is reflected by the share of fixations on that attribute:

$$y_{i0jt} = \bar{\beta}_{ijt} + \xi_{i0jt}, \tag{3}$$

where y_{i0jt} is the share of fixations on attribute j as percent of the total number of fixations for participant i at time t; $\bar{\beta}_{ijt}$ is the share of attention on attribute j for participant i at time t; and ξ_{i0jt} is unobserved heterogeneity.

In the absence of additional information about participant i or attribute j that could influence the share of attention on attribute j for participant i at time t, the model specifies:

$$\bar{\beta}_{ijt} = \bar{\beta}_{0jt} + \zeta_{ijt}, \tag{4}$$

where $\bar{\beta}_{0jt}$ is overall attention share on attribute j at time t; ζ_{ijt} is unobserved heterogeneity for participant i and attribute j at time t; $\zeta_{ijt} \sim N(0, \sigma_{jt}^2)$; σ_{jt}^2 is participant-level variance in attention shares for attribute j at time t.

Building on RIT (Caplin and Dean 2015; Matějka and McKay 2015), the model in this chapter assumes that consumers focus more on those brands that are worth the effort which implies that these brands receive more than their fair share of attention $(1/B)$. Specifically, the subjective value of brand b on attribute j for participant i at time t is a ratio between its share of attention and the expected share if all brands were equally preferred $(\theta_{ibjt} = \bar{\theta}_{ibjt}/B^{-1})$. The share of attention for participant i, attribute j, and brand b is reflected by the share of fixations on that brand and attribute relative to the total number of fixations on the attribute:

$$y_{ibjt} = \bar{\theta}_{ibjt} + \xi_{ibjt}, \tag{5}$$

where y_{ibjt} is the share of fixations on brand b and attribute j relative to the total number of fixations on attribute j for participant i at time t; $\bar{\theta}_{ibjt}$ is the subjective value of brand b and attribute j of participant i at time t; ξ_{ibjt} is unobserved heterogeneity.

If no other information is available about participant i, brand b, or attribute j that could influence the share of attention at time t, then:

$$\bar{\theta}_{ibjt} = \bar{\theta}_{0bjt} + \zeta_{ibjt}, \tag{6}$$

where $\bar{\theta}_{0bjt}$ is overall attention share on brand b and attribute j at time t; ζ_{ibjt} is unobserved heterogeneity for consumer i, brand b, and attribute j at time t; $\zeta_{ibjt} \sim N(0, \sigma^2_{bjt})$; σ^2_{bjt} is participant-level variance in attention shares for brand b and attribute j at time t.

The example in Table 4.2 illustrates the idea of the proposed model for one of the participants in the sample who chose brand C from the set of four TVs. The attribute fixation shares (y_{i0jt}) at the end of the decision process $(t = 3)$ are: 45%, 25%, and 29% on eco-, performance-, and other attributes respectively. Comparing these shares to the expected 30%, 40%, and 30% (based on the number of attributes of each type), indicates that: (1) the participant focused more on attributes that offer information about how eco-friendly the brands are $(\frac{45\%}{30\%} = 1.5)$, (2) to the detriment of performance attributes that are inspected less than expected $(\frac{25\%}{40\%} = .6)$, (3) while still fixating on other-attributes $(\frac{29\%}{30\%} = 1.0)$. The chosen brand (brand C) receives a similar number of fixations as compared to the other brands on eco-attributes, but a larger share on the other two attribute types. This suggests that the participant did not choose brand C based only on the information about its eco-attributes and formed an overall evaluation of the brand. The total number of fixations per brand could correctly predict brand C is likely to be chosen, but the attribute shares indicate why this brand was chosen over the other options on display.

Table 4.2 Fixations frequencies and shares for brands and attributes during choice (Participant 5, eco-goal and low time-pressure condition, brand C is chosen, TVs)

	Brand A	Brand B	Brand C	Brand D	Fixations on attribute	Attribute share (y_{ibj})
Eco-attributes	9 (19%)	13 (27%)	15 (31%)	11 (23%)	48	45%
Performance-attributes	1 (4%)	2 (7%)	20 (74%)	4 (15%)	27	25%
Other attributes	6 (19%)	4 (13%)	18 (58%)	3 (10%)	31	29%
Fixations on brand	16	19	53	18		

Note: numbers in brackets are shares of fixations for the attribute and brand combination relative to the total number of fixations on the attribute (y_{ibj}).

4.2.2 Effects of Decision Goals and Time Pressure on Attention

Equations 1-6 specify a general link between eye movements, attention, and brand utility that is agnostic to participant and brand characteristics. When data about such characteristics are available, the model can easily be adapted as we explain in this section.

In line with previous research on the impact of motivation and time pressure on eye movements during choice (Pieters and Warlop 1999), the model accounts for differences due to decision goals and time pressure:

$$\bar{\beta}_{ijt} = \eta_{0jt} + \eta_{1jt} X_i^{goal} + \eta_{2jt} X_i^{tp} + \eta_{3jt} X_i^{goal} X_i^{tp}, \quad (7)$$

where η_{0jt} is the importance of attribute j at time t for participants in the performance-goal and low time-pressure condition; η_{1jt} is the deviation from η_{0jt} for participants in the eco-goal and low time pressure condition; η_{2jt} is the deviation from η_{0jt} for participants in the performance-goal and high time pressure condition; and η_{3jt} is the deviation from the sum of η_{0jt}, η_{1jt}, and η_{2jt} for participants in the eco-goal and time-pressure condition.

$$\eta_{kjt} = \gamma_{kj1} + \gamma_{kj2} D_1 + \gamma_{kj3} D_2, \tag{8}$$

where D_1 and D_2 are dummy variables equal to 1 if t larger than or equal to 1 and 2 respectively; γ_{kj1} is the effect in the first time interval, γ_{kj2} is the deviation from γ_{kj1} in the second interval; γ_{kj3} is the deviation from the sum of γ_{kj1} and γ_{kj2} in the final interval.

Participants in the eco-goal condition ($X_i^{goal} = 1$) are expected to have larger attention shares ($\eta_{11} > 0$) for eco-attributes than participants in the performance-goal condition ($X_i^{goal} = 0$). Because participants under high time pressure are more likely to filter information, the share of attention to attributes that are most relevant to the decision goal is likely to increase. For example, participants in the eco-goal and high time-pressure condition should allocate a larger attention share on eco-attributes ($\eta_{31} > 0$) while participants in the performance-goal and high-time pressure condition should have a larger attention share on performance-attributes ($\eta_{22} > 0$).

The specification in equation 7 incorporates results of previous research, which pre-dates the current formalization in RIT, that finds consumers skip more information when they are under time pressure and that the type of skipped information is non-random (Pieters and Warlop 1999). These previous results describe the link between the total number of fixations and consumer-specific characteristics (task motivation and time-pressure). The proposed model extends these previous results in two ways: (1) it investigates the effects of consumer-specific characteristics on covert attention shares by accounting for measurement error in eye-fixations (ξ_{i0jt} in eq. 3) and (2) it models these effects over time as specified in eq. 8.

Between-consumer variation in decision goals and time pressure influence not only the attribute attention shares, but also the brand and attribute attention shares. For example, participants in the eco-goal condition ($X_i^{goal} = 1$) are expected to attend more to eco-attributes of the eco-friendly brand(s). These participants are more likely to discriminate more between the brands on the eco-attribute dimension than participants in the performance-goal condition. For the latter, differences between brands on the eco-attribute are not as relevant and hence not worth the effort to separate the brands on this dimension. High time pressure ($X^{tp} = 1$) should also influence differences between goal conditions (Bettman et al. 1998; Pieters and Warlop 1999) as participants who have less time to thoroughly inspect the brands benefit if they quickly identify the best brand on the most important attribute. Therefore:

$$\bar{\theta}_{ibjt} = \eta_{0bjt} + \eta_{1bjt}X_i^{goal} + \eta_{2bjt}X_i^{tp} + \eta_{3bjt}X_i^{goal}X_i^{tp}, \tag{9}$$

where η_{0bjt} is the average attention share for brand b and attribute j at time t for participants in the performance-goal and low time-pressure condition; η_{1bjt} is the deviation from η_{0bjt} for participants in the eco-goal and low time pressure condition; η_{2bjt} is the deviation from η_{0bjt} for participants who are in the performance-goal and high time pressure condition; and η_{3bjt} is the deviation from the sum of η_{0bjt}, η_{1bjt}, and η_{2bjt} for participants in the eco-goal and time-pressure condition.

$$\eta_{kbjt} = \gamma_{kbj1} + \gamma_{kbj2}D_1 + \gamma_{kbj3}D_2, \tag{10}$$

where D_1 and D_2 are dummy variables equal to 1 if t larger than or equal to 1 and 2 respectively; γ_{kbj1} is the effect in the first time interval, γ_{kbj2} is the deviation from γ_{kbj1} in the second interval; γ_{kbj3} is the deviation from the sum of γ_{kbj1} and γ_{kbj2} in the final interval.

4.2.3 Brand Choice Predictions and Model Estimation

If the decision goal of the participant is known, then predicting brand choice should be relatively easy. However, in a real-life choice setting decision goals are rarely known to the researcher. To test the extent to which eye movements reflect fundamental attention and utility accumulation processes, the proposed brand choice model (eq. 1-6, when the decision goal of the participant is unknown) is compared against five competing specifications. The six models differ in the amount and type of information they extract from eye movements (Table 4.3). The equations that describe the competing models are detailed in Appendix E.

Table 4.3 Brand choice predictions - model comparison

			Attribute importance revealed by		Subjective value		Hit rate			
	Brand Intercepts	Goal effects	Brand choice	Eye movements	Brand	Brand and attribute	Choice task			
	α_b	α_b^{goal}	β	β_{ij}	θ_{ib}	θ_{ibj}	1	2	3	4
M0	x						39%	31%	37%	42%
M1	x	x					51%	44%	59%	66%
M2	x		x		x		74%	58%	76%	73%
M3	x		x			x	75%	58%	73%	73%
M4					x		41%	30%	41%	40%
M5				x		x	67%	49%	70%	67%

The first two models incorporate standard choice models specification. M0 is a logit model that specifies utility as: $u_{ib} = \alpha_b + \varepsilon_{ib}$. The underlying assumption is that the observed component of brand utilities is the same for all participants (α_b), and that differences in brand choice between participants are due to unobserved utility shocks (ε_{ib}). M1 includes information about the decision goal of the participant and specifies: $u_{ib} = \alpha_b + \alpha_b^{goal} X_i^{goal} + \varepsilon_{ib}$. This implies that participants who have the same decision-goal evaluate the observed utility component ($\alpha_b + \alpha_b^{goal} X_i^{goal}$) in the same way, but that their choices differ due to idiosyncratic utility shocks (ε_{ib}). M0 and M1 do not include time-specific information, therefore these models do not update their brand choice predictions over time. Because M1 includes brand and decision-goal intercepts, it is expected to have a better hit rate than M0.

The next two models combine characteristics of the logit model (accounting for brand fixed intercepts) with measures extracted from eye movements. M2 specifies that brand utility is: $u_{ibt} = \alpha_b + \beta\theta_{ibt} + \varepsilon_{ibt}$, which combines the assumption of M4 + M0. The fourth model, M3, specifies brand utility as: $u_{ibt} = \alpha_b + \beta_j\theta_{ibjt} + \varepsilon_{ibt}$. By comparing M3 against M5 we test the extent to which attention-based importance weights (β_{ij}) are similar to choice-based attribute weights (β_j). Parameters α_b, β, and β_j are estimated conditional on observed brand choice. M2 and M3 can make moment-to-moment predictions because θ_{ibt} and θ_{ibjt} are time-specific and estimated based on observed eye movements up to time t.

The remaining two models predict brand choice using only attribute importance and subjective values extracted from eye movements. M4 is a naïve model that specifies brand utility as a function of subjective brand values (θ_{ibt}) only: $u_{ibt} = \theta_{ibt} + \varepsilon_{ibt}$. M5 is the proposed model (eq. 1-6). Because M4 and M5 extract the utility components (θ and β) from eye movements only, the choice probabilities can be calculated from moment-to-moment, before observing brand choice for any of the participants. By comparing M4 and M5 we test if the allocation of attention between the attributes provides additional insights into brand utilities in addition to the brand attention.

By comparing M5 and M0 we test the extent to which attention and utility are closely linked. M1 is akin to a segmentation approach because it estimates brand preferences within two separate groups of participants (eco-goal and performance goal). Subscript t is not included, because the logit models can only be estimated after observing brand choice.

Attention shares (eq. 3-6), effects of decision goals and time pressure on attention (eq. 7-10), and competing brand choice specifications (M0-M3) are estimated using MCMC techniques in R (R Core Team 2018; Stan Development Team 2018), using non-informative priors, and multiple chains with dispersed starting values. Convergence is assessed using potential scale reduction (Gelman and Rubin 1992).

To summarize, the model developed in this chapter extract participant and time-specific attribute importance weights (β_{ijt}) and subjective values of brand features (θ_{ibjt}) from eye-moments that participants make during choice. This is different from the standard approach which estimates attribute importance weights from brand choices that consumers make from sets in which attribute levels are systematically manipulated (e.g. conjoint design). The model uses differences in the amount and timing of eye movements between the brands and attributes to assess the effects of (1) time pressure and (2) decision goals on participants' preferences for attributes and brands.

4.3 Data

4.3.1 Participants and Design

Students at a large public university (*N* = 443) participated in a study on consumer preferences. They were informed that they would make choices between brands in four different product categories (light bulbs, travel mugs, TVs, and combi-fridges), on a simulated website, and that they would be given information about four brands. The data were collected in two waves: January-February 2016 and November 2016, in order to increase the number of participants. To motivate participants to make choices that align with their preferences, they were told that there would be a lottery after all participants complete the study. The prize of the lottery was one of the products that the winner chose during the experiment and this was an extra reward on top of the compensation they received that day in the lab. There were two lottery draws and two winners, one for each of the two data collection waves.

The experiment had a 2 (time-pressure) x 2 (decision goal) x 2 (brand order) between-subjects design and participants were randomly assigned to one of eight possible conditions. Time pressure conditions are high ("only 15 seconds") and low ("30 seconds"). Goal conditions are eco ("Choose the option that is most friendly to the environment") and performance ("Choose the option that has the highest performance"). Order conditions reverse the order of the brands on the display (ABCD and DCBA) to control for location effects on eye movements (Atalay et al. 2012). Participants were informed that for each of the choices they make, they would see three screens: (1) the description of the choice task, (2) the website with product information, and (3) the four brand names. Then, to help them familiarize with the task, participants had a practice round to choose one out of four brands of toothbrushes. Before every choice task, participants were reminded of the decision-goal and time-pressure condition they are in. The study ended with questions about how they have made the choices (e.g. what information they found important) and if they wanted to participate in the lottery. Those who indicated that they wanted to participate in the lottery wrote their email address and participant number on a lottery ticket which they placed in a "lottery" box on the research assistant's desk.

4.3.2 Eye-tracking Procedure

Participants were informed that their eye movements were going to be recorded. The experimental room had three cubicles each containing a Tobii TX300 eye-tracker arranged on a table and a chair in front of it. Participants were seated in front of the screen such that the centre of the screen was on the same level as their eyes. The distance between the eyes and the screen was approximately 60 centimetres. Then, the research assistant performed the calibration task. After calibration, the research assistant would leave the cubicle and the participant would start the study. All stimulus slides were projected on the screen and participants could see the next slide by making one click. The slides containing brand information were presented for a fixed amount of time (15s and 30s) and did not allow the participant to move to the next slide by clicking.

4.3.3 Eye-movement Measures

Tobii T60XL eye-trackers have a display resolution of 1920 x 1200 pixels and a sampling rate of 60Hz (Tobii AB 2015). Raw eye-movement data exported from Tobii Studio were grouped into eye-fixations using the Binocular-Individual Threshold algorithm (van der Lans et al. 2011). Then, we checked the percent of eye movements classified as fixations for all participant and task combinations (van der Lans and Wedel 2017) and retained $N = 334$ participants with at least 80% of their eye movements classified as eye-fixations for each of the four tasks ($n = 295$) or who have an average of at least 80% over all four choice tasks and a minimum of 70% in one of them ($n = 39$). There are $N = 334$ participants, $C = 4$ choice tasks, $T = 3$ intervals, $J = 3$ attribute-types ($j = 1$ (Eco), $j = 2$ (Performance) and $j = 3$ (Other)), and $B = 4$ brands. Hence, the total number of observations is 48,096. Table 4.4 presents raw data summaries grouped by goal and time pressure condition type.

Table 4.4 Raw data summaries

Condition	Time pressure	Decision Goal	N	Fixation durations (ms)		Number of fixations		Sum of fixation durations (s)	
				M	SD	M	SD	M	SD
1	Low	Eco	96	200.8	121.2	126.4	14.9	25.4	1.7
2	Low	Performance	79	199.6	119.2	127.4	14.8	25.4	1.8
3	High	Eco	87	195.2	103.5	66.3	8.5	12.9	0.7
4	High	Performance	72	191.2	93.6	67.9	6.3	13.0	0.7

Note: the difference between the sum of fixations durations and the time available to inspect the brands is due to saccades, and to fixations outside the AOI area.

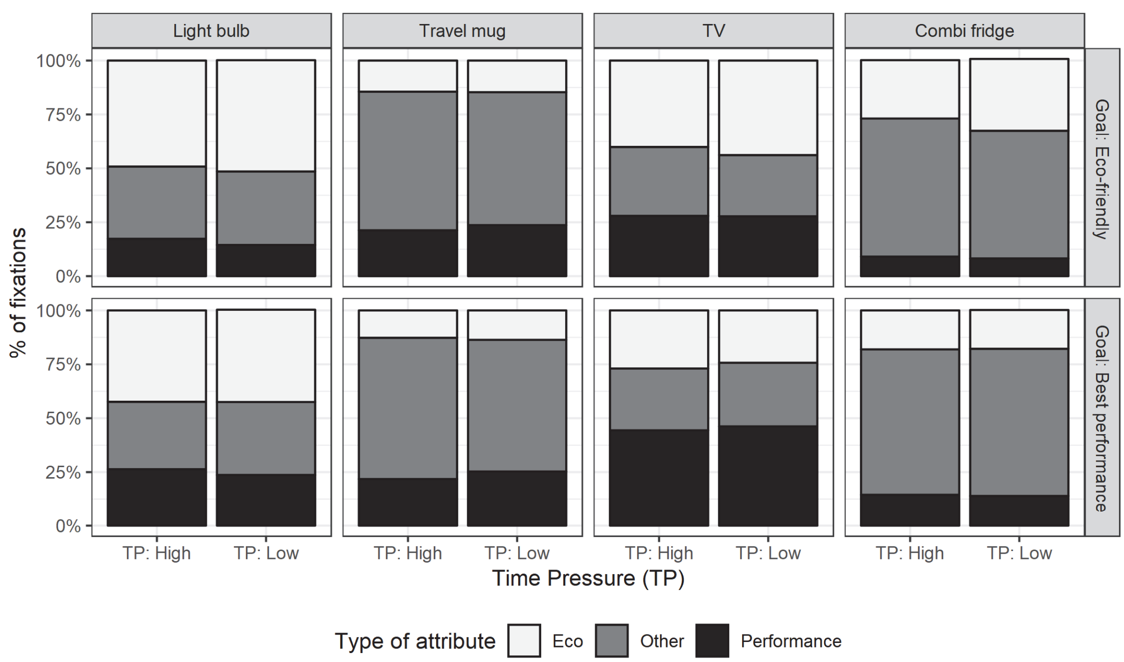
4.4 Results

In this section we first present model-free evidence based on: (1) observed differences in eye movements across participants, brand, attributes, and time, and (2) differences in brand choice between participants. Then, the estimated effects of decision goals and time-pressure on attention during choice between TV brands are discussed in 4.4.2 and 4.4.3 respectively. To save space, the estimated effects for the other three product categories are included in Appendix F. In section 4.4.4 we describe the results of brand choice predictions for the six model specifications introduced in section 4.2.3.

4.4.1 Model-free Evidence

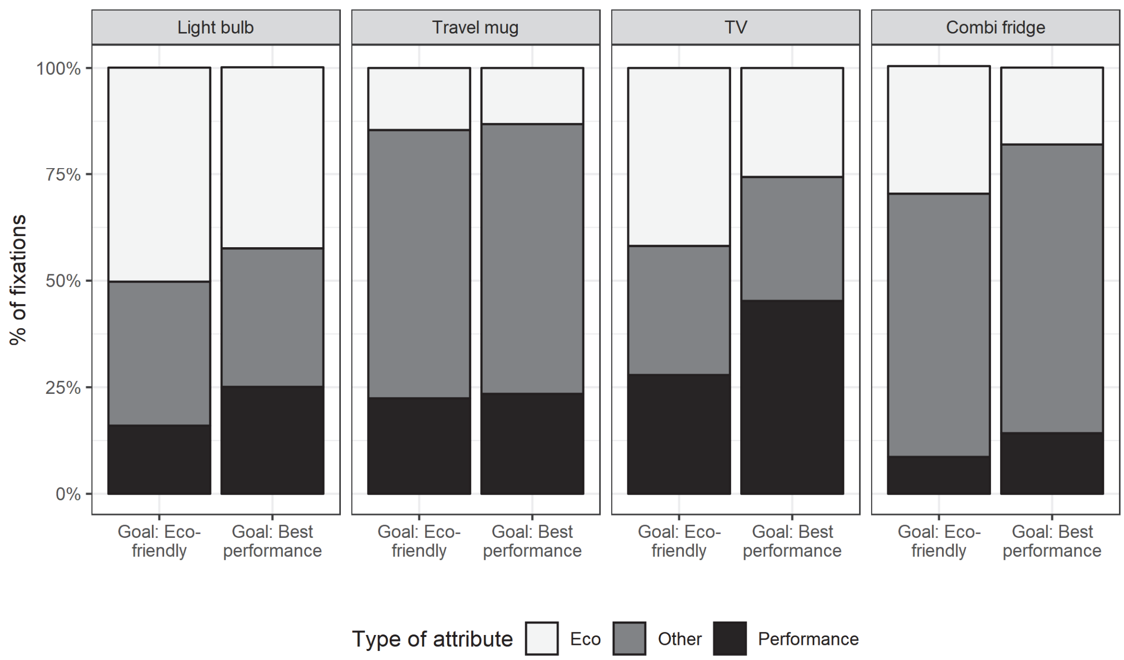
Figure 4.2 shows the average fixation shares for the three attribute types in each choice task, by decision-goal and time-pressure condition. It appears that there are only very small differences due to time-pressure. Fixation shares by attribute type differ between tasks because the number of attributes of a type varies between the product categories (Table 4.1).

Figure 4.2 Minor differences in shares of fixations by attribute type between participants in high and low time-pressure conditions



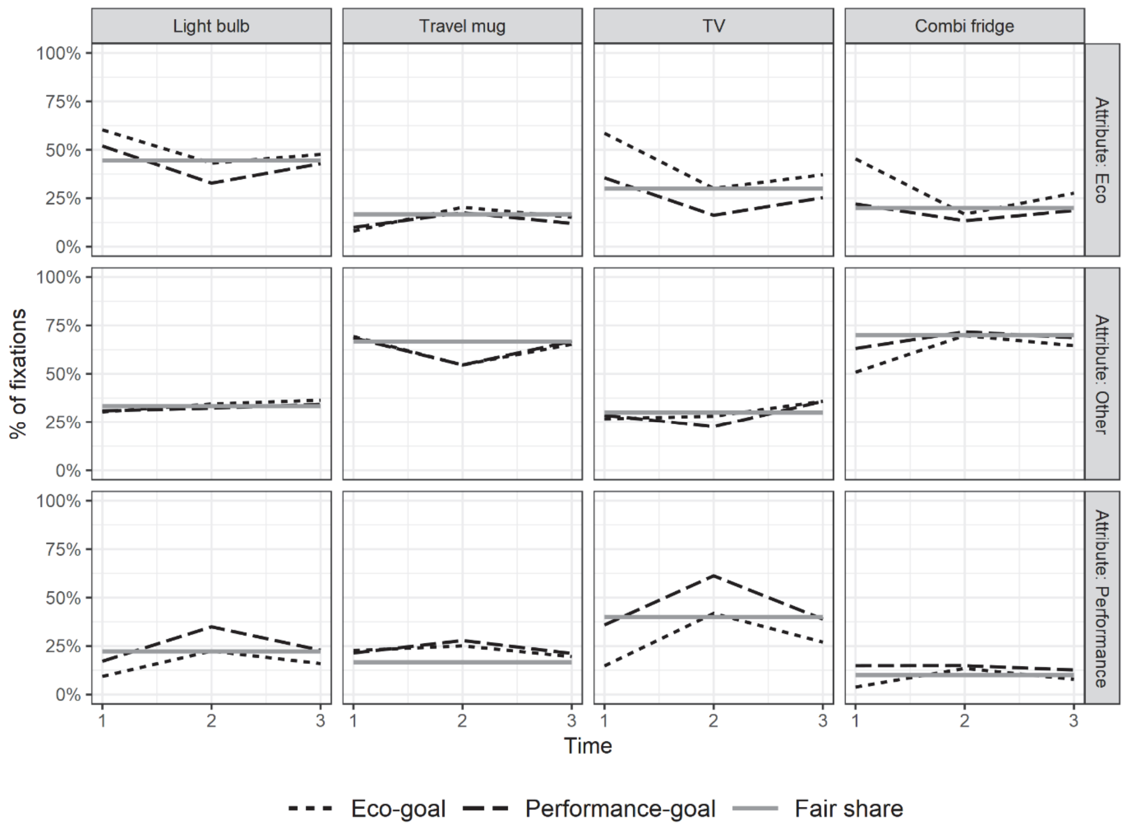
Participants in the eco-goal condition allocate a larger share of fixations to eco-attributes. For three of the four choices (light bulb, TV, and combi-fridge) this difference is easily noticeable, while for the remaining choice task (travel mug) this difference is much smaller (Figure 4.3). The larger share of fixations on eco-attributes for participants in the eco-goal condition reduces the share of fixations on performance-attributes, while the share of fixations on other attributes is similar between participants in the two decision goal conditions.

Figure 4.3 Overall, participants allocate larger fixations shares to the attribute type that is aligned with their decision goal



Participants in the eco-goal start by allocating a larger share of fixations on eco-attributes as compared to the fair-share (Figure 4.4, light bulb, TV, combi fridge tasks). In the second interval, these participants switch their focus to performance-attributes. In the final moments before expressing brand choice participants (both eco-goal and performance-goal groups) switch back to the attribute that matches their decision goal. For two of the four tasks (light bulb and TV), participants in the performance-goal start by having a larger fixation share on eco-attributes, as compared to the fair-share. This could be explained by the fact that eco-attributes include pictorial information (EU energy efficiency labels), which usually attracts more attention due to visual salience effects (Pieters and Wedel 2004; van der Lans et al. 2008b). Three of the product categories (light bulb, TV, combi fridge) include these labels, in line with EU regulations.

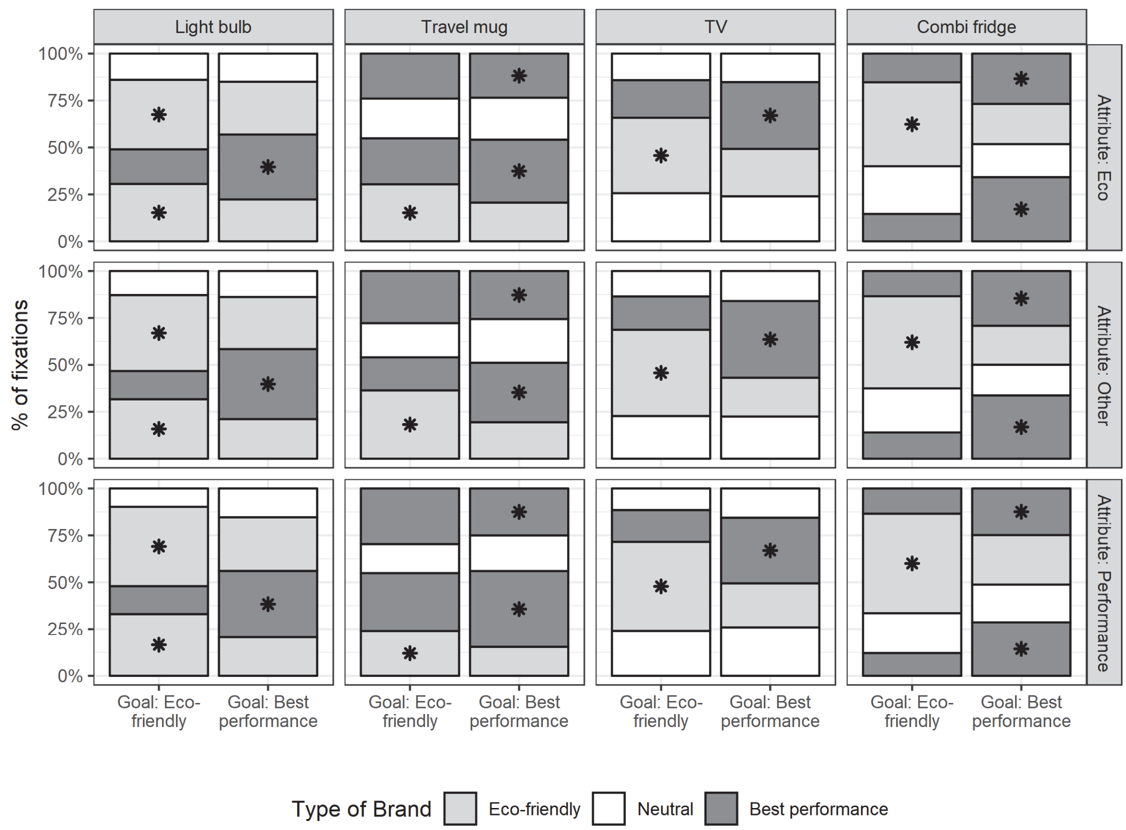
Figure 4.4 From moment-to-moment, participants allocate more fixations to the attribute type that is most relevant for their decision goal



Note: shares are calculated using the number of fixations within each time interval.

We now present evidence for differences in fixation shares between brands. Brands that are normatively the best option have larger fixations shares on all attribute types (Figure 4.5). For example, eco-friendly brands for participants following an eco-goal and the brands with best performance for participants in the performance-goal group. Brand C (light bulb) is normatively the best option for participants in the performance-goal group. This brand receives a larger fixation share for each of the attributes for participants in the performance-goal group as compared to participants in the eco-goal group. Similar differences are observed also for the other brands and choice tasks.

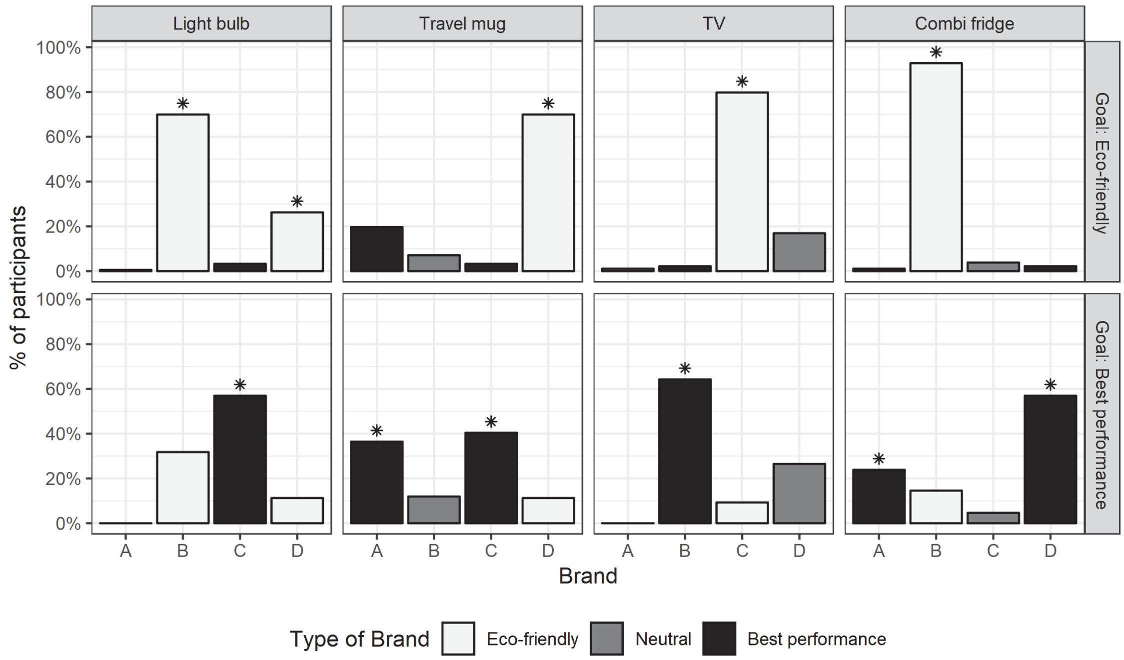
Figure 4.5 Participants fixate more on the attributes of the brand that matches their goal (*)



Note: * indicates normatively best option: Eco-friendly brand for participants in the eco-goal group and the brand with best performance for participants in the performance-goal group.

In addition to fixating more on attributes and brands that match their goal, participants also chose the brand that matches the decision-goal condition they were randomly assigned to (Figure 4.6). Participants instructed to choose an environmentally friendly option chose the brands that are eco-friendly (85% over all choice tasks). Similarly, participants in the performance-goal condition chose the matching brand (70% over all choice tasks).

Figure 4.6 Participants choose the brand that matches their goal



Note: * indicates normatively best option: Eco-friendly brand for participants in the eco-goal group and the brand with best performance for participants in the performance-goal group.

4.4.2 Attribute Attention Shares

We estimate the effects of decision goals and time pressure on attribute attention shares, as deviations from the fair share based on the number of attributes grouped within a type (Table 4.1). Results are consistent with previous studies that examine effects of time pressure and task motivation on attention during choice tasks (Pieters and Warlop 1999). We describe the results for choice between TV brands (results for the other choice tasks are in Appendix F).

In line with the model-free evidence in Figure 4.4, participants in the performance-goal and low time-pressure condition ("Performance-goal" in Table 4.5) allocate initially more than the fair share of attention to eco-attributes (.06, p-value < .001). During the second interval the importance of eco-attributes decreases (-.17, p-value <.001) while the importance of performance-attributes increases (.18, p-value <.001). In the final third, the importance of performance-attributes decreases to the initial level (-.19, p-value <.001) while the share of other-attributes (.12, p-value <.001) and that of eco-attributes (.07, p-value = .01) increase as compared to the second time interval. Participants in the performance-goal and high time-pressure condition attend more to performance-attributes in the second interval (.15, p-value <.001) as compared to participants in the performance-goal and low time-pressure condition.

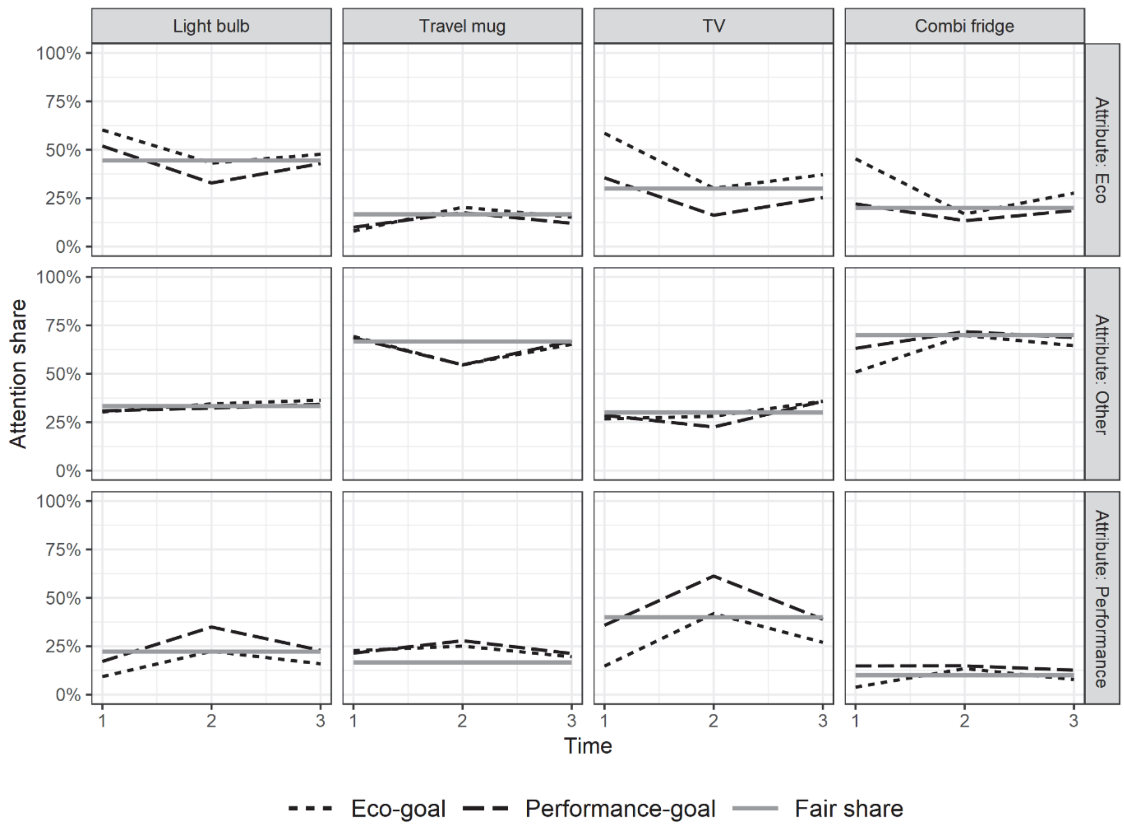
Table 4.5 Estimation results – attribute importance

	Period 1 (γ_{kj1})			Period 2 (γ_{kj2})			Period 3 (γ_{kj3})		
Choice task 3 (TV)	*M*	*SD*	p-value	*M*	*SD*	p-value	*M*	*SD*	p-value
Performance-goal ($k = 0$, eq. 7-8)									
Eco-attr.	.06	.02	.01	-.17	.03	<.001	.07	.03	.01
Performance-attr.	-.01	.02	.53	.18	.03	<.001	-.19	.03	<.001
Other attr.	-.05	.02	.05	-.01	.03	.56	.12	.03	<.001
Eco-goal ($k = 1$, eq. 7-8)									
Eco-attr.	.20	.03	<.001	-.10	.04	.02	.01	.04	.24
Performance-attr.	-.22	.03	<.001	.07	.04	.05	.02	.04	.23
Other attr.	.02	.03	.17	.03	.04	.18	-.03	.04	.46
High time-pressure ($k = 2$, eq. 7-8)									
Eco-attr.	.00	.03	.72	-.06	.05	.21	.04	.05	.14
Performance-attr.	-.06	.03	.06	.15	.05	<.001	-.06	.05	.17
Other attr.	.06	.03	.03	-.10	.05	.03	.02	.05	.21
Eco-goal x High time-pressure ($k = 3$, eq. 7-8)									
Eco-attr.	.07	.05	.048	.02	.06	.24	-.06	.06	.30
Performance-attr.	.01	.05	.24	-.11	.06	.10	.12	.06	.03
Other attr.	-.08	.05	.07	.09	.06	.07	-.06	.06	.32

Note: attr = attributes. Estimates in the "Period 2" column indicate changes in attribute attention shares relative to the "Period 1" column. Estimates in the "Period 3" column indicate changes in attribute attention shares relative to the "Period 1" and "Period 2" columns. Shares within a condition and time interval sum to 0. The values are as compared to the "fair share" based on the number of attributes of each type as compared to the total number of attributes. See table 4.1.

Participants in the eco-goal and low time-pressure condition start by allocating a larger share of attention to eco-attributes (.20, p-value < .001), taking away from the share of performance attributes (-.22, p-value < .001), as compared to participants in the performance-goal and low time-pressure condition. The importance of eco-attributes for participants in the eco-goal and high time-pressure condition is initially larger than that for participants in the eco-goal and low-time pressure condition (.07, p-value = .048).

Figure 4.7 Attribute types aligned with participants' decision goals receive more attention – changes over time in estimated attribute shares



To summarize, participants allocate larger attention shares to attribute types that match their decision goal. This happens already in the first interval for participants in the eco-goal condition (.20, p-value <.001) with an additional increase if they are under high time-pressure (.07, p-value = .048). This initial advantage diminished over time though. In the second interval, participants in the eco-goal condition lower the share on eco-attributes an additional 10 percentage points (.10, p-value = .02) as compared to those in the performance-goal and low time-pressure group. This can be due to eco-attributes being more visually salient (EU eco-labels). Participants in the performance-goal fixate more on performance-attribute in the second interval (.18, *p*-value <.001), with an additional 15 percentage points if they are under time-pressure (.15, *p*-value <.001). We plot the estimated attribute attention shares, which capture similar trends to those in fixations shares (Figure 4.7 compared to Figure 4.4)

4.4.3 Brand-and-attribute Attention Shares

Table 4.6 presents the effects of decision goal and time-pressure conditions on the brand attention shares within an attribute type. These are deviations for the fair share of .25 (there are four brands). Shares should sum to 0 within a condition, attribute type, and time interval. When shares do not sum to zero it is because some participants do not have fixations on any of the brands for the respective attribute-type within the time interval (e.g. performance-attributes for participants in the eco-goal and low time-pressure condition). For example, participants in the eco-goal and low time-pressure condition allocate a smaller share of attention to performance attributes (-.22, *p*-value < .001). Based on the raw number of fixations, 23% of participants in the eco-goal and low time-pressure condition do not fixate on performance-attributes during the first time interval.

The eco-attributes of the eco brand in choice task 3 (brand C) attract more attention (than the .25 fair share) only from participants in the eco-goal both initially (.11, *p*-value <.001) and in the second interval (.15, *p*-value <.001). This brand attracts more attention also on the other two attributes from participants in the eco-goal. When these participants have more time to inspect the brands (low time-pressure), already in the first interval they inspect brand C more than the other brands for performance (.17, *p*-value <.001) and other attributes (.17, *p*-value <.001). Participants who follow an eco-goal but are under high time-pressure inspect performance attributes of brand C in the second interval (.13, *p*-value = .03) and other attributes in the third interval (.15, *p*-value = .02). This shows that under increased time-pressure participants focus first on the more relevant attributes in order to identify the brand that matches their goal. Then, to the extent that they have time left, they inspect the other information available and form an overall evaluation of the brand they are about to choose.

Participants in the performance-goal allocate more attention to the matching brand (B) in the second interval both for performance attributes (.09, *p*-value < .001), eco-attributes (.15, *p*-value < .001) and other attributes (.15, *p*-value < .001). Participants under high time-pressure start from a lower attention share on performance attribute (-.09, *p*-value = .01).

Table 4.6 Estimation results – subjective values of brand-and-attribute combinations

Choice task 3 (TV)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Performance-goal ($k = 0$, eq. 9-10)										
Eco-attributes	A	-.03	.02	.25	-.08	.04	.03	-.05	.04	.18
	B (Perf)	.03	.03	.12	.15	.04	<.001	.00	.04	.25
	C (Eco)	.01	.03	.22	-.08	.04	.02	.00	.04	.25
	D	-.01	.03	.52	-.08	.04	.047	.05	.04	.07
Performance-attributes	A	-.08	.03	<.001	-.01	.04	.61	-.05	.04	.17
	B (Perf)	.02	.03	.16	.09	.04	<.001	.08	.04	.01
	C (Eco)	-.06	.03	.02	.03	.04	.13	-.05	.04	.16
	D	-.03	.03	.17	.04	.04	.14	-.01	.04	.69
Other-attributes	A	-.02	.03	.43	-.12	.04	<.001	.01	.04	.23
	B (Perf)	.04	.03	.06	.15	.04	<.001	.04	.04	.11
	C (Eco)	.00	.03	.67	-.07	.04	.05	.00	.04	.25
	D	-.05	.03	.04	.02	.04	.22	-.01	.04	.68
Eco-goal ($k = 1$, eq. 9-10)										
Eco-attributes	A	-.04	.03	.21	.01	.05	.24	.05	.05	.12
	B (Perf)	-.05	.04	.13	-.20	.05	<.001	.01	.05	.25
	C (Eco)	.11	.03	<.001	.15	.05	<.001	.00	.05	.72
	D	-.02	.03	.53	.09	.05	.04	-.04	.05	.34
Performance-attributes	A	-.07	.03	.04	.03	.05	.18	.06	.05	.10
	B (Perf)	-.14	.03	<.001	-.05	.05	.30	-.11	.05	.03
	C (Eco)	.17	.03	<.001	.04	.05	.16	-.01	.05	.65
	D	-.07	.04	.03	.04	.05	.17	.03	.05	.21
Other-attributes	A	-.06	.03	.07	.08	.05	.048	-.03	.05	.46
	B (Perf)	-.10	.03	.01	-.18	.05	<.001	-.02	.05	.57
	C (Eco)	.17	.03	<.001	.17	.05	<.001	-.07	.05	.12
	D	.01	.03	.22	-.03	.05	.43	.07	.05	.06

(table continues on next page)

(table continues from previous page)

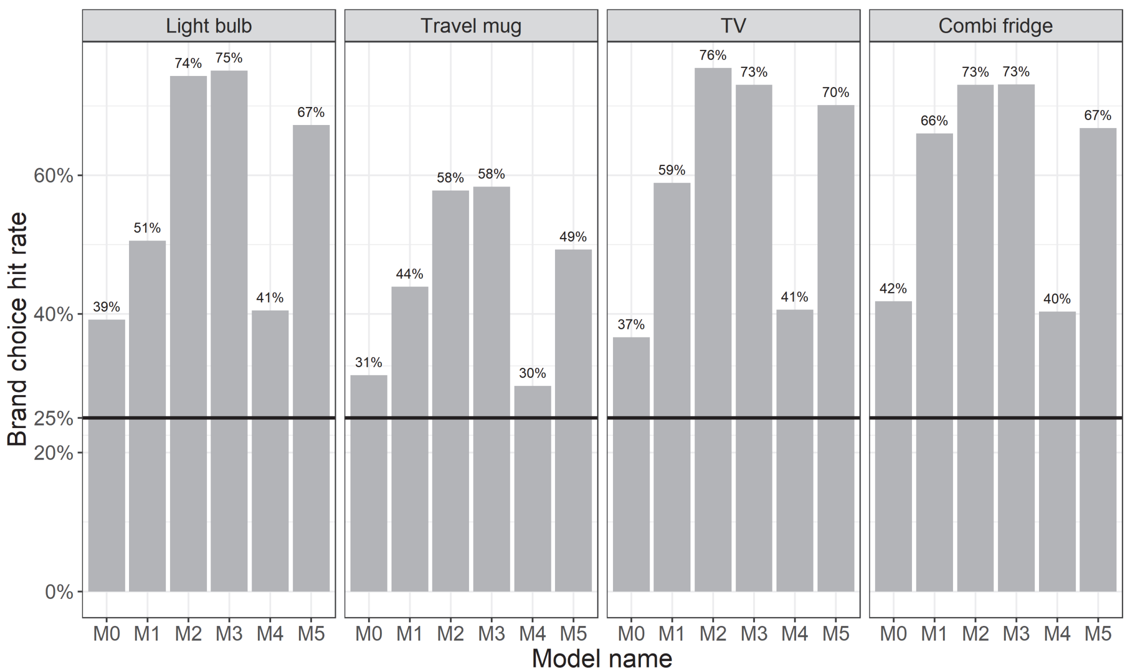
Choice task 3 (TV)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
High time-pressure ($k = 2$, eq. 9-10)										
Eco-attributes	A	.02	.04	.20	-.12	.05	.03	.06	.05	.10
	B (Perf)	-.08	.04	.047	-.13	.05	.01	.15	.05	.002
	C (Eco)	-.01	.04	.65	-.03	.05	.46	.06	.05	.12
	D	-.02	.04	.51	-.03	.05	.47	.01	.05	.24
Performance-attributes	A	-.01	.04	.63	-.01	.05	.71	.00	.05	.25
	B (Perf)	-.09	.04	.01	.03	.05	.21	-.03	.05	.43
	C (Eco)	-.04	.04	.30	.05	.05	.15	.02	.05	.23
	D	.00	.04	.73	.03	.05	.21	.01	.05	.24
Other-attributes	A	.02	.04	.20	-.02	.05	.59	.00	.05	.71
	B (Perf)	-.07	.04	.06	-.01	.05	.68	.04	.05	.17
	C (Eco)	.00	.04	.72	-.02	.05	.59	.01	.05	.24
	D	.00	.04	.25	-.04	.05	.42	.03	.05	.19
Eco-goal x High time-pressure ($k = 3$, eq. 9-10)										
Eco-attributes	A	-.07	.05	.13	.18	.07	.004	-.06	.07	.35
	B (Perf)	.05	.05	.13	.14	.07	.03	-.15	.07	.04
	C (Eco)	.00	.05	.25	-.06	.07	.34	-.02	.07	.60
	D	.10	.05	.02	-.09	.07	.22	-.02	.07	.64
Performance-attributes	A	-.04	.05	.35	.01	.07	.25	.01	.07	.25
	B (Perf)	.05	.05	.12	-.05	.07	.45	.08	.07	.12
	C (Eco)	-.13	.05	.01	.13	.07	.03	-.03	.07	.56
	D	.00	.05	.74	-.02	.07	.60	-.05	.07	.43
Other-attributes	A	-.04	.05	.36	.04	.07	.21	.00	.07	.74
	B (Perf)	.08	.05	.06	-.02	.07	.63	-.04	.07	.52
	C (Eco)	-.06	.05	.24	.00	.07	.25	.15	.07	.02
	D	.03	.05	.19	.06	.07	.17	-.17	.07	.01

Note: Shares should sum to 0 within a condition, attribute type, and time interval. When shares do not sum to zero it is because some participants do not have fixations on any of the brands for the respective attribute-type within the time interval (e.g. performance-attributes for participants in the eco-goal and low time-pressure condition). For example, participants in the eco-goal and low time-pressure condition allocate a smaller share of attention to performance attributes (-.22, p-value < .001). Based on the raw number of fixations, 23% of participants in the eco-goal and low time-pressure condition do not fixate on performance-attributes during the first period. Estimates in the "Period 2" column indicate changes in brand-and-attribute attention shares relative to the "Period 1" column. Estimates in the "Period 3" column indicate changes in brand-and-attribute attention shares relative to the "Period 1" and "Period 2" columns. The values are as compared to the "fair share" of .25.

4.4.4 Brand Choice Predictions

The proposed model (eq. 1-6) specifies that the accumulation of brand utilities is a function of attribute importance weights (β_{ijt}) and participants' subjective values of the brand-and-attribute combinations (θ_{ibjt}). These utility components (β_{ijt} and θ_{ibjt}) are extracted from observed eye movements that participants make during choice. Of the six specifications for brand utility and brand choice (introduced in section 4.2.3 and summarized in Table 4.3), four of them include time specific information and therefore can be used to predict brand choice from moment to moment. The remaining two (M0 and M1) specify brand utility as a function of brand and participant information that does not vary over time. We start by presenting the brand choice hit rate for the six models, at the moment of choice, so after all eye movements are observed (Figure 4.8).

Figure 4.8 Brand choice hit rate at the moment of choice



Note: M0-M3 predict brand choice based on different combinations of attention-based and choice-based estimation results (Table 4.3); M4 predicts brand choice based on attention-based subjective values for each brand; M5 predicts brand choice based on attention-based attribute importance and subjective values for brand and attributes (eq. 1). The black line marks the 25% random hit rate if the four brands on display were equally likely to be chosen.

The first group of models (M0 and M1) specify that attention does not provide any information about brand utility. This is a core assumption of standard choice models that assume consumers use all the available information. M0 predicts brand choice based on estimated brand specific utility intercepts[7]. The hit rate of this model is similar to that of the naïve M4. For two of the tasks M4 has a better hit rate (2 percentage points for task 1 and 4 percentage for task 3) while for the other two tasks M0 performs slightly better (1 percentage point for task 2 and 2 percentage points for task 4). If the decision-goal of the participant is known, then including this information in the model (M1) improves hit rate, though not enough to outperform the proposed model (M5). M5 has a larger hit rate for all product categories as compared to M1 (respective percentage points increase:16, 5, 11, and 1).

The difference between M0 and M1 shows that knowing the decision goal of a participant offers the opportunity to predict their choice to the extent that they are decision-goal compliant. However, this is rarely the case in real-life consumer choice. We argue that the difference in hit rate between M1 and M5 is evidence for the fundamental link between eye movements, attention, and utility accumulation processes.

M4 and M5 predict brand choice using only information extracted from eye movements. These models are completely agnostic to the specific brand information on the display. For example, M4 uses the attention allocated to each brand, but the model does not know the brand type (eco- or performance-goal matching) or the decision goal of the participant. M5 is the model proposed in equations 1-6. By comparing M4 and M5 we see that decomposing brand utilities into attribute importance and subjective brand-and-attribute value leads to a large improvement in brand choice hit rate (between 19 and 29 percentage points over the four choice tasks). Table 4.3 include the hit rate for all the four choice tasks. This shows that in the absence of any other information about the brands (e.g. brand type, market shares) or the participant (e.g. decision goal), a model that discriminates between attribute types is better able to capture brand utilities and to predict brand choice.

---

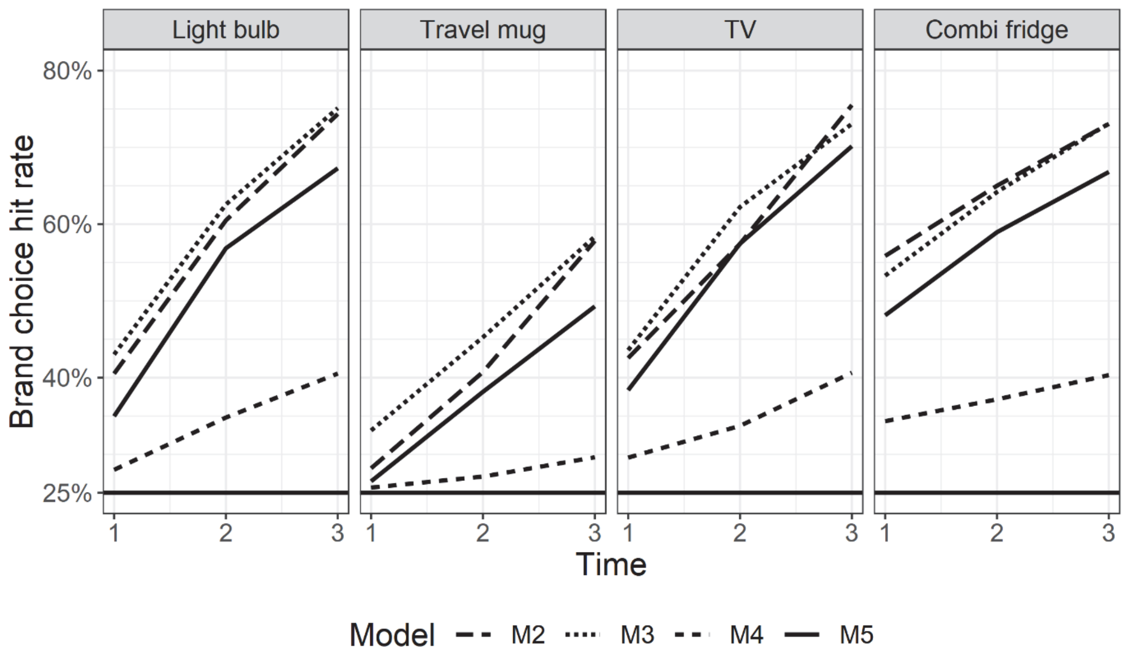
[7] As an additional check, we estimate a model without brand fixed effects, but with partworths for attribute levels. The hit rates of this model are the same as that of M0.

The second group of models combines characteristics of the previous two groups. M2 combines assumptions of M4 and M0 by specifying that brand utility is a function of brand intercepts and brand specific subjective values extracted from eye movements. The link between the subjective value of each brand and brand utility is estimated based on the observed brand choice. Different from M5, M3 estimates the link between brand-and-attribute subjective values and brand utility conditional on observed brand choice. The difference in hit rate between M5 and M3 is due to: (1) including brand intercepts (M3), and (2) estimating attribute importance based on observed choice (M3) vs. extracting attribute importance for observed eye movements (M5). M2 and M3 have similar hit rates – the only differences are for task 1 (1 percentage point in favor of M3) and task 3 (3 percentage points in favor of M3). Based on the hit rate at the moment of choice, after all eye movements are observed, it seems that separately extracting brand and attribute specific subjective values is not necessary (M2 and M3). Accounting for differences in attention between brands per attribute type is more important when we make predictions based only on attention estimates (M5 vs M4). However, the conclusion should be more nuanced given the differences in hit rate over time (Figure 4.9) which we discuss next.

While M2 and M3 both improve over M5 it is important to note that the models use different amounts of information. Both M2 and M3 use the observed brand choice to estimate the link between attention and utility and to predict brand choice. M5 makes brand choice predictions using only attention measures extracted from eye movements.

Four of the six models include time-varying information in the specification of brand utility and therefore can make moment-to-moment predictions before brand choice is observed. Figure 4.9 presents these results compared to a naive 25% hit rate (there are four brands in the set). For all choice tasks the models predict above the 25% threshold already in the first interval. The performance is better for choice tasks 1, 3, and 4 which are also the tasks which larger differences in attention shares between participants in the two goal conditions (Figures 4.7 and 4.4). The hit rate improves over time for all the four models, with larger changes for M5, M2, and M3. M5, the model that predicts brand choice based only on attention measures improves over time at a similar rate as M2 and M3. This suggests that including brand intercepts and thus accounting for brand choice shares is important in setting the start point of the hit rate.

Figure 4.9 Brand choice hit rate over time



While M2 and M3 have very close hit rates at the moment of choice, over time these differences vary. For example, for TVs, M2 has a hit rate of 76% at the end, 3 percentage points larger than that of M3, but during the previous time intervals M3 had a better hit rate: 5 percentage points at time 2 and 1 percentage point at time 1. This is also the case for task 2: 4 percentage points at time 2 and 5 at time 2.

4.5 Discussion

This chapter aimed to test whether eye movements reflect both what brand consumers are going to choose, and why this brand is preferred. By manipulating decision goals in a between-subjects experimental design, we were able to examine (1) how eye-movement patterns reflect otherwise unobserved attention processes closely related to by participants' decision goals, and (2) how attribute importance and subjective brand-and-attribute values extracted from eye movements are able to capture the utility accumulation processes that take place during value-based choice. The proposed model extracts attribute importance and subjective brand-and-attribute values from how consumers allocate their attention over time between brands and attributes. We now summarize the most important results. Then, we discuss the implications that these results have for related theories of attention and choice, and end with limitations and opportunities for future research.

First, we find that attribute importance shares change over time. Initially, participants are more likely to focus on the attribute that matches their goal. For participants following the eco-goal, these are eco-attributes (e.g. energy consumption), while for participants in the performance goal, these are category specific attributes that indicate the quality of the brand (e.g. sound and image quality for TVs). Then, they inspect other attributes and towards the end of the process go back and focus again on the attributes that match their goal. This supports the idea that the contribution of brand attention to utility is time-varying, which we first introduced in chapter 2.

Second, participants quickly discover the brand that matches their decision goal. However, they continue to inspect other attributes of this brand and form an overall evaluation of its utility match. Even though participants were assigned decision goals that could easily be implemented by simply choosing the best brand on the attribute of interest (eco or performance), they still inspected the other information about the brands.

Third, attention-based importance weights and subjective brand values capture brand utilities as indicated by the results of the brand choice predictions (M5). This happens already in the first third of the decision time and then improves as more eye movements are observed. By the time all eye movements are observed, the proposed model has an average brand choice hit rate of 63%, 8 percentage points larger than that of a logit model that includes information about the type of brand that the consumers aim to choose (55%).

Fourth, the results suggest that including brand utility intercepts and estimating attribute importance weights from choice (M3 vs M5) improves brand choice predictions. At the moment of choice, this is an improvement of 8 percentage points on average: from 63% for M5 to 71% for M3 (one brand from a set of four).

4.5.1 Implications

Just as in previous chapters, the model developed in this chapter builds on RIT (Caplin and Dean 2015) and SSM (Forstmann et al. 2016). While specific applications of SSM make different assumptions about what exactly is accumulated over time (evidence in aDDM (Krajbich et al. 2010) or valence in MDFT (Roe et al. 2001)), they make similar assumptions about the more general idea that the utilities of the brands on display evolve over time as the consumer inspects them. In that sense, the proposed model is similar – we specify that utilities evolve over time as the consumer inspects the brands, similar to both aDDM and MDFT. Then, building on MDFT the model accounts for attribute importance weights and subjective value of brand-and-attribute combinations. The proposed model differs from models of MDFT in two important ways. First, we model brand specific utilities instead of the relative advantage of each brand as compared to the other options (valence in MDFT (Roe et al. 2001)). Second, the model developed in this chapter infers participant-specific subjective values for brand-and-attribute combinations from sequences of eye movements. While the general MDFT model specifies participant-specific subjective values for brand-and-attribute combinations (Roe et al. 2001), empirical applications make two simplifying assumptions (Diederich 2003; Dror et al. 1999). First, that participants derive the same values from the different attribute levels. Second, that these values are equal to the corresponding attribute levels, which requires all attribute levels to be numeric.

We find that participants allocate different amounts of attention not only between the attributes, but also within an attribute over the different brands. Such differences in attention between brands at the level of an attribute challenge an important assumption of MDFT. Specifically, that once attention is focused on an attribute, all the brands in the set are evaluated on that dimension. The implication of these results is better aligned with the general assumption of RIT - that not all brands are evaluated on all attributes (Matějka and McKay 2015).

Decision field theory was originally developed for decision making under uncertainty (Busemeyer 1985; Busemeyer and Townsend 1993). Thus far, MDFT models adapted to value-based choice have mostly been compared to other decision models (e.g. elimination by aspects, weighted additive utility model) based on their ability to capture and explain context effects (similarity, attraction, compromise) (Berkowitsch et al. 2014; Roe et al. 2001). The results of this chapter indicate that models in this literature could be further adapted, by for example using attention-based measures of subjective values for brand-and-attribute combinations. This would make the model more generally applicable to choice situations since it would eliminate the need for multiple choices within the same brand category. Because the proposed estimation algorithm can easily be implemented in open-source software (Stan Development Team 2018), the new model would be more accessible to decision researchers who want to implement it. The standard MDFT is difficult to estimate, as it requires numerical integration techniques (Berkowitsch et al. 2014), since there is no closed form solution for the choice probabilities (Roe et al. 2001).

The second implication for MDFT is related to attribute importance weights. In our model, attribute importance is extracted from observed eye-fixations shares, after accounting for measurement error and unobserved sources of heterogeneity at participant level, such as different decision strategies (Shi et al. 2013) or reliance on different types of information (van der Lans et al. 2008a). In MDFT models, these weights are assumed to change from moment to moment stochastically. However, it is not clear what drives this stochastic process and how decision goals could be implemented. We speculate that specifying the structure of what influences importance weights would improve the model by allowing its parameters to be directly impacted by different sources of variation at the participant and/or attribute level.

The predictive performance of the proposed model (M5) as compared to standard models of choice (M0 and M1) suggests that attention-based subjective values of brand-and-attribute combinations are closely related to the utility accumulation process that takes place during choice.

The results of our model show that it captures a general link between eye movements, attention, and brand utility, which implies that it can predict choice for new consumers, in new categories, and for whom decision goals are not known. This would contribute to previous research that has focused on settings where the utility of the choice alternatives is measured from repeated choices or ratings. Our paper extends this by inferring utility from sequences of eye movements in settings where no previous choices and no prior information about preferences are available at individual level. More specifically, we focus on single choices that consumers make for complex brands, as is usually the case in product categories for expensive, durable goods (e.g. digital cameras). Unlike more standard methods of attribute importance measurement, the model extracts participant, brand, attribute, and time specific attribute importance weights and subjective brand values from eye movements that participants make during a single choice.

The importance weights and subjective values can be influenced both by characteristics of the decision context that are outside the control of the consumer (e.g. decision time restrictions) and by internal decision goals. The proposed model quantifies the effects of these variables on changes in importance weights and subjective values over time. This extends previous models that extract time-invariant importance weights from how a participant uncovers specific information on an IDB (information display board) while the value of the brands/items is equal between participants (Wedell and Senter 1997).

4.5.2 Limitations

In the introduction to this chapter, we argue that our proposed method is preferred to traditional methods for assessing attribute importance because it does not require multiple choices per participant. At the same time, because we calibrate the proposed model on single brand choice, it is not possible to directly compare our results with those of other methods (e.g. conjoint design). Future research could focus on this and examine if attention-based (the model developed in this chapter) and choice-based (conjoint design) attribute importance and subjective values of brand-and-attributes are similar. If differences between the two methods appear, then it would be interesting to explore what drives them (e.g. are the methods measuring different constructs?).

While attention measures allow us to predict brand choice ahead of time, and the results demonstrate a strong link between eye movements and choice, we cannot claim this is a causal link from eye movements to brand utilities (or in the other direction). The assumption in our model is that attribute importance is reflected in how participants inspect the brands and attributes, in line with prior research (Wedell and Senter 1997). However, we cannot exclude the possibility that in other decision tasks, when participants are not as motivated to follow a certain goal, more salient information might attract attention and potentially influence choice. This could be partly influenced by visual salience, the extent to which some attributes or brands stand out due to the perceptual features of the area on the display (color, luminance, edges) (van der Lans et al. 2008b). But it could be also salience driven by the extent to which an attribute or brand stands out as compared to the average characteristics of the alternatives in the choice set, or the extent to which attributes or brands depart from a reference good (Bordalo et al. 2016). Future research could investigate under what conditions attribute importance weights are mainly driven by visual salience, decision goals, or characteristics of the alternatives in the choice set.

The fair share of attribute attention used to calculate attribute importance weights is based on the number of attributes, and thus assumes that all attributes have similar levels of complexity or require the same number of eye-fixations to be inspected. We argue that in the study used in this chapter, the complexity was similar between attributes, but when this is not the case, the fair share can be calculated based on a different rule (e.g. area on the screen, number of words, reading time). This would change the effects of the estimated effects for participants in the performance-goal and low time-pressure condition, but not the effects for the other conditions, which are relative to this baseline.

Chapter 5

Conclusions and Onwards

5.1 Introduction

Consumers frequently make brand choices online after inspecting multiple alternatives. Hence, by the time consumers purchase a brand, they have made numerous other choices: what information to inspect, for which alternatives, for how long, and at what moments. All these choices reflect how consumers allocate attention and time resources between the different brands and attributes over time. The three empirical essays in this dissertation focus on specific aspects of this type of brand choice processes. In this chapter, we start by summarizing the results of the studies included in this dissertation (section 5.2). Then, we discuss implications of these results for theories of decision making (section 5.3), marketing managers (section 5.4), and consumer protection policies (section 5.5). The chapter and dissertation conclude with suggestions for next steps towards a theory of rational attention (section 5.6). While the results of this dissertation show that attention reveals and predicts choice, we do not claim that this is proof of 'rational attention'. We acknowledge that much remains to be done and introduce four propositions for a theory of rational attention that provide directions for future research.

5.2 Summary of Main Results

The models developed in this dissertation are calibrated on eye tracking and brand choice data in order to gain insight into the moment-to-moment utility accumulation that occurs before choice is expressed. Table 5.1 presents an overview of the studies. The three empirical essays investigate the link between attention and utility and the extent to which eye movements, as manifest indicators of attention (Just and Carpenter 1980; Liechty, Pieters, and Wedel 2003), can predict *why* consumers choose a brand, *what* brand they choose, and *when* they express their choice. At the same time, each of the essays aims to answer specific research questions. Chapter 2 focuses on *what* brand consumers choose, chapter 3 on *what* brand and *when* consumers choose, and chapter 4 develops a model that examines *why* consumers choose a certain brand. In this section we present the main findings in each of these empirical essays (summarized in Table 5.2).

Table 5.1 Overview of studies

Chapter	Participants (sample size)	Product category* (#brands on screen)	Decision time	Participant characteristics	Brand characteristics
2	US consumers (N = 342)	Smartphones (B = 5)	Freely determined by participants	Information complexity (manipulated**) Smartphone ownership	Brand ownership, brand name
3	NL Students (N = 214)	Digital cameras (B = 4)	Freely determined by participants	Decision goal (manipulated**)	Position of brand on the screen
4	NL Students (N = 334)	Light bulbs (B = 4) Travel mugs (B = 4) TVs (B = 4) Fridges (B = 4)	Externally controlled	Decision goal (manipulated**) Time pressure (manipulated**)	Attributes (number of attributes varies between 6 and 10)

Note: *participants make only one choice per product category; **between-subjects design.

5.2.1 Chapter 2

When information is available at the same time on display, consumers are free to inspect it as they see fit (e.g. by brand, by attribute). The pattern in which information is inspected reflects underlying cognitive processes that take place during choice (Bettman et al. 1998; Shi et al. 2013). Because these processes evolve over the course of the decision, using eye movements that reflect them offers the possibility to answer two research questions: (1) how trajectories of attention to each of the brands during the choice task contribute to the accumulation of

Table 5.2 Summary of key findings

Chapter	Focus	Attention	Outcome	Predictions	Key findings
2	What fundamental attention processes reflect brand utility	Attention quantity and three types of attention (integration, comparison, other)	*What* brand is chosen	Brand choice (in sample fit)	Attention for integration is better able to reflect utility (as compared to other types of attention). The link between attention and utility is stable across different information density displays. Substantial heterogeneity in attention between participants and brands, even when accounting for specific effects (e.g. loyalty).
3	Within choice dynamics: Attention reflects how two types of utility (brand and search) change over time until choice is expressed	Amount and timing of attention for integration.	*What* brand is chosen and *when* brand choice is expressed	Brand choice and moment of choice (from moment-to-moment, K-fold cross validation)	Very early in the process (3rd decile), attention predicts what brand is going to be chosen and the moment when this choice is expressed. Participant heterogeneity in decision thresholds.
4	Allocation of attention between attributes reflects their importance. Allocation of attention between brands reflects how specific attribute levels are valued.	Attribute attention shares. Brand attention shares (within an attribute).	*Why* a brand is chosen (which attributes and brand characteristics contribute to brand utility)	Brand choice (out of sample predictions based on attention only)	Eye movements reflect which attributes are more important for brand choice. Attention during choice predicts what brand is going to be chosen for new participants and in new product categories.

utility and final choice, and (2) which fundamental attention processes contribute to the accumulation of utility and brand choice. To answer these questions, we develop a model rooted in recent developments in cognitive science and economics (Caplin and Dean 2015; Krajbich and Rangel 2011; Reutskaja et al. 2011). It (1) infers the trajectories of four types of attention from observed eye-movement measures, and (2) quantifies the link between the attention trajectories and brand utilities.

Prior research on attention and choice (Krajbich et al. 2010; Pieters and Warlop 1999; Yang et al. 2015) has often focused on one type of eye movements (i.e. fixations) while largely ignoring eye-saccades, which hold the promise of providing information about the specific attentional processes that people are engaged in (Rayner 1978; Shi et al. 2013). The essay proposes and tests the following assumptions. First, we argue that (1) consumers engage in rapid brand comparison and information integration processes, which are reflected in, respectively, between-brands saccades and within-brand saccades, and that (2) in particular the latter predict choice. Second, the model investigates the extent to which prior ownership effects on choice operate via attention.

The model is calibrated on data collected from consumers ($N = 342$) who had expressed to be in the market for a smartphone, randomly drawn from large, locally representative participant pools, from three locations in the continental US. The results show that attention for integration (reflected by eye-saccades within the same brand) contributes significantly to brand utility, even when attention quantity (reflected in the number of eye-fixations on the brand) is already accounted for. In addition, we find support for the time-varying contribution of attention to brand utility. This provides initial evidence for the link between attention and brand choice.

The results of this chapter show substantial heterogeneity in attention patterns over time between consumers and brands. The models developed in chapters 3 and 4 build on this and account for unobserved sources of heterogeneity in attention (ϵ_{ibt}) and eye movements (ξ_{ibt}). Importantly, when consumers are presented more information, they adapt how much attention they invest in the task, but not how this is allocated over time between the brands. This is supported by results that show differences in information density impact the amount of attention allocated during choice, but not the brand-specific brand trajectories.

5.2.2 Chapter 3

Consumers choose not only which brands to inspect and in what sequence, but also for how long to inspect the brands before choosing one of them. Models of RIT abstract from the timing of attention within one choice (Steiner et al. 2017). While the aDDM allows the amount of time a brand is under focus to influence choice (*what* and *when*), it assumes that eye-fixations are random with respect to the value of the fixated item (Krajbich et al. 2010). We argue that next to the amount of attention allocated to a brand, the moments when the brand is under focus reveal how brands accumulate utility until a threshold is reached. The proposed model uses moment-to-moment brand attention to predict both what brand consumers are going to choose and when they express this choice.

The model developed in chapter 3 builds on the results of chapter 2. First, it uses eye-movement measures that combine characteristics of attention quantity and attention for integration. Specifically, it consolidates consecutive eye-fixations within the same brand, separated by at most one fixation on another brand. Then, the model extracts brand-and-consumer specific attention from these measures that retain information about the number of eye-fixations (attention quantity) and the brand that they belong to (attention for integration). It does so while accounting for measurement error and consumer- and brand-heterogeneities in attention over time.

Specific to chapter 3 is the focus on both brand choice and moment of choice. The proposed model makes predictions of what brand consumers are going to choose and at what moment in the future based on sequentially observed eye movements up to the moment when the prediction is made. This is important because it provides insights into within choice dynamics as compared to existing research on dynamic decision making in marketing, which primarily focuses on dynamics over repeated choices (for a review of the literature, see Kumar and Man Luo 2008). While applications of SSM, such as the aDDM (Krajbich et al. 2010; Krajbich and Rangel 2011), explicitly focus on computational processes that take place during a choice, the basic nature of the tasks and the required prior preference measurement limit the applicability of these models to consumer choice from complex brands.

While the results of chapter 2 suggest that brand choice can be predicted based on intermediate attention measures, these predictions can only be made once eye movements up to the moment of choice are observed. Hence, the hit rates in chapter 2 are closer to a measure of in-sample hit. Chapter 3 addresses this limitation and develops a more general method that predicts brand choice and moment of choice based only on eye movements observed until the time when predictions are made. As more eye movements are observed, the predictions are updated. Competing models are compared based on their prediction performance assessed using a K-fold cross validation approach. The results show that brand choice and moment of choice can already be predicted above change after only 30% of the decision time (seven brand visits given the average decision time of 22 brand visits).

An important result of Chapter 3 is that during choice, consumers adjust the decision threshold that brands need to reach before they are chosen. This provides insights into consumer heterogeneity in decision time. More importantly, it is a necessary first step towards influencing consumer choice in real time. Specifically, in order to optimize actions meant to influence consumer choice, it is necessary to predict not only what brand it more likely to be chosen, but also how much time there is to intervene.

5.2.3 Chapter 4

In chapter 2 brand utilities are a function of different attention types and the link between attention and utility is estimated from choice data. In chapter 3 we use a similar approach but extend the model to accommodate moment-to-moment predictions of brand choice and moment of choice. But still, the link between attention and utility is estimated from the observed choices of participants in a calibration sample. Both chapter 2 and 3 specify brand utility as a function of attention at the brand level. Chapter 4 extends this and develops a model that specifies brand utility as a weighted sum of the consumer's subjective values for the features of each brand. Importantly, both the weights (attribute importance) and the subjective values of the brand-and-attribute combinations are extracted from observed eye movements. We test the extent to which attention-based attribute importance weights capture consumer preferences as compared models that estimates these weights conditional on observed choice. The hit rate of the proposed model outperforms that of standard choice models that use information about the type of brand participants aim to choose, which in a more realistic setting is unlikely to be known before any choice is observed.

The model developed in chapter 4 extends previous studies that assume heterogeneity only in importance weights, but not in the value assigned to the features of the choice alternatives (Wedell and Senter 1997) and that measure consumer's subjective values (1) from preference ratings, independent of attention (Krajbich et al. 2012) or (2) from repeated choices in a conjoint design (Yang, Toubia, and De Jong 2015).

Specifically, the proposed model examines how eye movements during choice reflect attribute importance weights that (1) are closely linked to consumers' decision goals and that (2) influence the utility of the brands on display.

5.3 Implications for Research on Value-based Choice

The first chapter of the dissertation introduces two literature streams that have focused on how "limits on attention impact choice" (RIT) (Caplin and Dean 2015, p. 2183), and on how evidence for choice alternatives accumulates during choice (SSM) (Krajbich et al. 2010; Ratcliff et al. 2016; Roe et al. 2001). Theories of rational inattention (Caplin and Dean 2015) and sequential sampling models (Busemeyer 1985) highlight the important role of attention during choice as a selection mechanism and as modulator of fixed brand preferences. The three empirical essays of this dissertation build on this theoretical foundation of fundamental processes that take place during choice. However, we argue that the role of attention during choice is not limited to filtering out information (RIT) or compensating for differences in otherwise fixed brand values (aDDM). Therefore, in addition to *what* brands consumers attend to and *for how long*, the sequence of moments *when* consumers focus on the brands plays an important role as well. This positions attention as a coordinating mechanism through which consumers implement their decision goals - an overarching idea for the models developed in this dissertation. The results of this dissertation offer new insights into fundamental processes of attention and utility accumulation and thus contribute to a new perspective of the role of eye movements in consumer choice. Because attention plays such an important and complex role during choice, it has important implications for theories of decision making, which we discuss in this section.

SSM and RIT models explicitly integrate (in)attention as a core component. In other literature streams, the role of attention is implied but not explicitly modeled. Some of these previous models offer competing explanations for the effects that we find, as well as inspiration for measures of uncertainty that could improve the models developed in this dissertation. We start by discussing these and then present implications for SSM and RIT.

5.3.1 Uncertainty during Brand Choice

The model developed in Chapter 3 specifies that uncertainty influences the utility of continuing to inspect the brands. Specifically, uncertainty is operationalized using an entropy measure that captures the extent to which brands have different choice probabilities. The results do not find support for a significant effect of uncertainty on search utility and implicitly on decision time. We argue that there is merit in investigating these effects further and provide alternative measures for the four types of uncertainty introduced in Chapter 1.

Type I: Uncertainty about the description of a brand on a specific attribute. A measure that captures this type of uncertainty needs to be at the brand and attribute level. One study that accounts for this type of uncertainty combines repeated choice data (choice-based conjoint (CBC) design) with eye movements that participants made during the tasks (Yang et al. 2015). This model specifies that consumers need several fixations on a specific cell in order to acquire the brand and attribute information it contains (Yang et al. 2015, p. 170, Eq. 1). Before any fixations, the probability that the consumer knows the information in the cell depends on the number of levels of that attribute ($^1/_{L_j}$, where L_j is the number of levels for attribute j). The results for this specific study show that after only one fixation on a cell, the probability that participants acquired the corresponding information about laptop attributes (e.g. CPU, price, hard drive) is .87, while after a second fixation the probability is more than .99. Even though this model was calibrated using up to 16 choices per participant, the parameter that captures the amount of information extracted in one fixation is identified only at the aggregate level. However, because participants in CBC studies make repeated choice between different combinations of a limited set of attribute levels[8], it is possible that from one choice to another they learn what levels to expect. This is supported by results that show that

[8] Six attributes with four levels each (Yang et al. 2015); six attributes with two to four levels (Meißner et al. 2016).

information search decreases as participants progress through the choice tasks (Meißner et al. 2016; Yang et al. 2015). Hence, such an approach needs to be adapted if differences in information acquisition between consumers, attributes, or brands are expected. For instance, if consumers have experience with some of the brands or if purchases within the category are frequent, then the measure would need to account for memory effects. Similarly, if there are differences in complexity between attributes (e.g. rating stars are easier to understand as compared to more technical information), these should also be included in the model.

Type II: Uncertainty about the importance of an attribute can be difficult to separately identify from type I uncertainty. Even though learning about attributes and learning about preferences are different at a conceptual level, they are equivalent from a modelling perspective as long as brand utilities are linear in attributes and preference weights (Ching et al. 2013). Due to the mediating role of attention for the effect of advertising on attribute importance (MacKenzie 1986), accounting for consumers' past exposure to advertising is one direction that future research can explore. While advertising has been shown to influence brand choice, these results are based on market-level measures of advertising expenditure (dollar amount, advertising frequency) (Bronnenberg and Dubé 2017). Hence, they do not account for consumer-level heterogeneity in advertising exposure. The implicit assumption that consumers within the same market of geographic area have been exposed to the same advertisements becomes increasingly difficult to justify given the recent growth in consumer-level targeting. Online platforms have detailed information about what advertisements a particular consumer has been exposed to and engaged with. Because these platforms have data about exposure but also other actions that consumers have taken to reduce uncertainty (e.g. watching product reviews on YouTube), these types of data can be used to get a more accurate measurement of advertising effects (inform consumers, influence preference formation). Then, they could optimize how information is displayed to account for the link between past attention to advertising and attribute importance.

In experimental studies, one idea could be to add a description of the attributes and their levels prior to the choice task. This would make sure that participants acquire information about the possible attributes levels before seeing the brand combinations on display. Then, a smaller proportion of the eye movements during the task would be needed for reducing type I uncertainty. However, it is difficult to claim that consumers would not already start learning about their preferences, especially if the attributes have different numbers of levels. If consumers were to follow a two-step decision process, then they would be more likely to process by attribute: first acquire the information, then decide out how important that attribute is. However, there is evidence of the contrary – consumers frequently switch between processing by brand and processing by attribute (Shi et al. 2013). In addition, a two-stage approach implies that it would not be possible to predict brand choice as early as we do in the three essays of this dissertation (after 25-35% of the decision time).

Type III: Uncertainty about the overall utility of a brand. Brand utilities are updated from moment to moment as the consumer inspects the information on display. This offers the possibility to capture uncertainty as a function of changes in brand utility over time. If these changes are frequent, relatively large, and both positive and negative, then this would suggest that the consumer is unsure. If the consumer continues to inspect the brand while these changes are ongoing, then the expectation of that brand's utility is probably high enough to justify the effort. The trade-off between inspecting brands that appear promising in order to resolve uncertainty about their utility and choosing the best alternative thus far is related to exploration vs. exploitation in the multi-armed bandit (MAB) literature (Gittins 1979) as well as models of sequential search and choice (SSC) (Weitzman 1979). In MAB problems, the objective of the consumer is to maximize the rewards from exploration (learning about the choice alternatives) and exploitation (making a final choice for one of the alternatives) (Meyer and Shi 1995; Powell and Ryzhov 2012). The objective in SSC models is similar: maximize the difference between the utility of the chosen brand and the costs of finding it.

In order to use the general idea of exploration vs. exploitation from the MAB literature in the type of decision processes that this thesis focusses on, a measure for the benefit of reducing uncertainty about the utility of a brand is needed. The reason is that learning about the utility of a brand is the reward that consumers receive from inspecting a choice alternative that they do not choose later on. One solution could be to measure the reduction in uncertainty and then estimate the benefit. This approach could provide additional insights into consumer characteristics that favour exploration vs. exploitation, to the extent that it is possible to empirically identify these effects (e.g. motivation, involvement in the category, forward looking behaviour, risk aversion). One measure that captures changes in probability distributions is the Kullback-Leibler divergence:

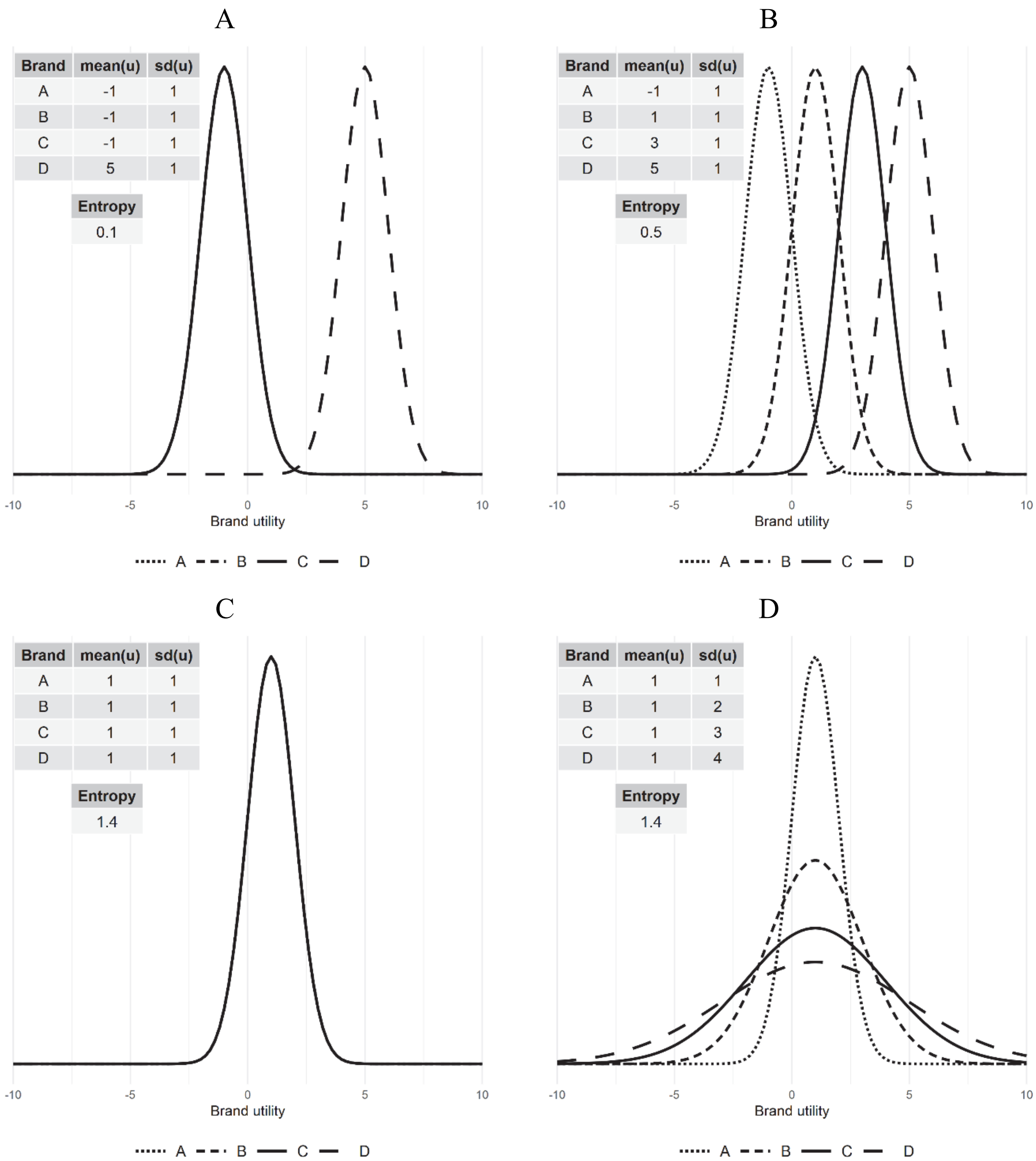
$$D_{it}^{KL}(t \parallel t-\Delta) = \iint p_{ibt}(u_{ibt}) \log \frac{p_{ibt}(u_{ibt})}{p_{ibt-\Delta}(u_{ibt-\Delta})} dud\Delta. \tag{1}$$

Where $D_{it}^{KL}(t \parallel t-\Delta)$ is the Kullback–Leibler divergence between the brand choice probability distribution at time t and $t-\Delta$; Δ is the time interval over which the change in choice probabilities is calculated (e.g. when $\Delta = 1$ $D_{it}^{KL}(t \parallel t-1)$ captures the change in brand choice probabilities from one moment to the next); u_{ibt} is the utility of brand b at time t for consumer i, p_{ibt} is the probability that consumer i chooses brand b at time t.

RIT models use a similar measure to capture changes in beliefs, with one important difference. RIT models use the expected change in choice probabilities over the remaining decision time(Steiner et al. 2017), while the measure in Eq. 1 integrates over past changes. These correspond to the expected gain from continuing to inspect the brands and the information gain that the consumer experienced thus far in the choice.

Type IV: Uncertainty about which brand has the highest utility in the set. In Chapter 3 the Shannon entropy measure captures the extent to which brands have similar choice probabilities. Shannon entropy is defined for discrete distributions and hence it does not account for the uncertainty around each brand utility (Figure 5.1 C and D). In addition, there are two other limitations to consider. First, it uses all brand choice probabilities, which is a limitation if consumers aim to differentiate only the chosen brand from the second preferred

Figure 5.1 Examples of changes in Shannon entropy as a result of changes in brand utilities



Note: Entropy values calculated as in Eq. 11 (Chapter 3) and rounded to one decimal. The minimum entropy is 0 when one brand has choice probability of 1. The maximum is 1.4 when brands have equal choice probabilities.

option. An alternative approach would be to include a measure of the distance between the first and second brands. Second, changes in the ranking of the brands do not always lead to changes in entropy. If the consumer represented in Figure 5.1B updates the utilities of brands C and D to 5 and 3 units respectively, then entropy remains 0.5. However, this change in the ranking of the brands would suggest that the consumer is unsure which of them is best. Ideally, a measure on uncertainty about which brand has the highest utility is sensitive to such changes.

5.3.2 Sequential Sampling Models (SSM)

The results of the empirical essays in this dissertation show that using attention, as compared to observed eye movements, is better able to capture the moment to moment utility accumulation that takes place during choice. Thus far, SSM applications that integrate eye movements use raw frequencies or durations of eye movements (Krajbich et al. 2010; Krajbich and Rangel 2011; Reutskaja et al. 2011). This comes with limitations. Using observed eye movements can lead to an underestimation of the link between attention and brand utility (chapter 2), with negative consequences for the accuracy of moment of choice predictions (chapter 3). Therefore, we argue in favor of specifying both search and brand utilities as functions of consumer, brand, and time specific attention extracted from eye-movement sequences.

The models we propose infer consumer's subjective values from attention and then estimate the link to utility (drift rate) conditional on choice. The specification can easily be adapted to any number of brands and attributes, which could be useful for extending the aDDM to more than the current limited number of options (3 brands with one attribute, or 1 brand with 2 attributes) (Krajbich et al. 2010; Krajbich and Rangel 2011).

The results of the dissertation show that consumers use eye movements to inspect those brands and attributes that are most relevant for their choice and that are more likely to be chosen. This contradicts assumptions of aDDM that fixations are random with respect to the value of the fixated item (Krajbich et al. 2010) and MDFT models that assume attention weights are identically and independently distributed over time (e.g. according to a Bernoulli process) (Roe et al. 2001).

5.3.3 Rational Inattention Theory (RIT)

The models in this dissertation build on RIT. While developments in this stream of literature recognize the important role of attention during choice, they argue that attention is essentially unobservable, just as preferences (Caplin and Dean 2015). Therefore, the proposed analytical models of RIT argue that both preferences and attention (or lack thereof) need to be recovered from observed choices (Masatlioglu, Nakajima, and Ozbay 2012). Specifically, from state dependent stochastic choice data (SDSC, introduced in section 1.3.2) collected by exposing participants repeatedly to decision problems that vary in terms of available actions, the value of the correct choice, and the prior probability of each state. In a marketing context, this implies that participants would make repeated choices from sets of brands whose attributes are systematically manipulated in a within-subjects experimental design, akin to choice-based conjoint studies. The models in this dissertation extract attention from eye-movement data, thus eliminating the need to observe multiple choices per participant. The characteristics of the data used to calibrate the proposed models in this dissertation differ from those of SDSC data. However, both this dissertation and RIT models are essentially interested in understanding the fundamental link between attention, utility, and choice. Therefore, the findings of this dissertation have implications for RIT models.

First, the results of chapter 2 indicate that (1) the moments when brands are attended to provide important information about changes in attention over time, and that (2) trajectories of attention, rather than the overall amount of attention, are more closely related to brand utilities (chapter 2). Therefore, we propose that changes in attention over time are accounted for, as they reflect not only what brands were under focus, but also for how long, and during which moments. Current analytical models of RIT account for what choice alternatives are attended to and the amount of attention (Matějka and McKay 2015), but assume that the timing of attention is irrelevant and can therefore be abstracted from (Steiner et al. 2017).

Second, the model proposed in chapter 3 accounts for the possible effects of uncertainty on the duration of the choice task. Specifically, participants who are not sure which of the brands provides most utility benefit from inspecting the brands for a longer time. This integrates a core assumption of RIT: that the costs of inspecting the brands is proportional to the choice-uncertainty of the decision maker, which is captured by an entropy-measure (Steiner et al. 2017). However, the results of the model comparison in chapter 3 (Table 3.2, and section 3.5.1) do not support this assumption. We argue that the general idea that uncertainty influences decision time and allocation of attention has merit and should be pursued in future research, as we discuss in more detail when we introduce the propositions for TRA (section 5.6). At the same time, we speculate that the entropy-based measure of uncertainty might not be best at capturing the underlying construct.

In information theory, Shannon entropy is a measure of uncertainty about a random variable. In RIT, this entropy measure captures the uncertainty of the decision maker with respect to the utility derived from choosing an alternative. In order to reduce uncertainty, the DM uses attention to select information that is relevant and filter out information that is not. Then, entropy-based costs imply that “the cost of multiple signals spread over many periods is identical to the cost of a single signal conveying the same information” (Steiner et al. 2017, p. 523). The assumption is that attention works like a filter and the consumer decides which “channel” to tune in and acquire information from. Then, the cost of the acquired information is not influenced by how easy or difficult it is to process it, but by the extent to which it makes the consumer more confident in selecting a brand.

5.4 Managerial Implications

“Knowing what to sell, when to sell, and to whom is essential for a firm to allocate scarce marketing resources in an efficient and effective manner” (Kumar and Man Luo 2008, p. 63). While eye-tracking data is not *yet* routinely collected, companies can use other types of data that reflect attention and that are easily available such as browsing history, geo-location, or purchase history. We argue that the fundamental link between attention and choice is present regardless of the manifest attention indicators (eye movements or browsing history). Online retailers (e.g. Amazon) and advertising platforms (e.g. Google Ads, Facebook) are already capitalizing on the ability to infer interests, attitudes, and purchase probabilities from detailed information about their users. Such platforms have access to large amounts of consumer level data that cover browsing behavior not only on their own website, but also on other websites though cookies. This data is valuable both for attribution models that quantify the contribution of every touchpoint prior to a final purchase (Danaher and van Heerde 2018) and for recommender systems.

The first essay focuses on online decisions in information-rich environments. These decisions are increasingly common as consumers have easy and fast access to detailed information. Websites display information on tens of attributes and consumers are expected to change their processing strategy in order to adapt to the amount of information (e.g. processing more by attribute). The findings of this essay suggest that consumers only change how much time and attention they allocate to the brands on display, but they do not change their strategy. This has implications for how online retailers display information about complex brands and suggests that they need to make sure the most diagnostic or important attributes are accessible to the consumer with a minimum level of effort.

The second essay focuses even more on the dynamic aspect of decision making. Theories on information processing stages and choice argue that consumers make decisions by going through a sequence of stages – information search, utility formation, and only later on choice. Under this framework, utilities are formed very late in the process, this process is unobservable, and preferences can only be inferred based on observed choices. This essay provides support for an early start of the utility formation process and for visual attention reflecting the moment-to-moment build up to choice. This can be used to improve preference measurement in situations when choices cannot be observed or there is reason to believe the observed choices are not an accurate reflection of consumers' true preferences. For example, in situations when consumers have an incentive to be dishonest (e.g. choosing an insurance plan in line with one's risky behaviors) or feel pressured to give a certain answer (e.g. attitudes towards substance use, unhealthy lifestyle choices, political preferences).

In order to make good product recommendations, online retailers need to have a measure of the utility match between the consumer and the recommendation. While the exact rules that power recommender systems are to a large extent unavailable to the public, some information is available in reports (Smith and Linden 2017) and patents (Lam et al. 2012). These offer information about how different companies implement variants of the two main approaches for recommendations: collaborative filtering or content-based filtering (Adomavicius and Tuzhilin 2005; Lam et al. 2012), which we discuss in more detail in sections 5.4.1 and 5.4.2 respectively.

E-commerce is the fastest growing market at the moment and this trend is expected to continue. In the US, online sales have a double-digit growth over the past years[9] and the leading online retailer Amazon has an active customer base of 300 million worldwide[10].

---

[9] https://www.census.gov/data/tables/2017/econ/arts/supplemental-ecommerce.html
[10] https://www.emarketer.com/Chart/US-Amazon-Active-Customers-vs-Prime-Members-2010-2019-millions/187753, last accessed on July 13, 2019

Consumers frequently use similar websites that display information on multiple brands and attributes in response to a search query, to help them choose products. The decisions made on these sites are often very fast, taken in just a few minutes or seconds (Shi et al. 2013) and both the websites and retailers featuring their products have great interest in understanding how such decisions are made. This dissertation looks into how consumers make such rapid decisions and examines additional factors that can play a role in these processes, such as information complexity and brand ownership. While the empirical results in the three essays provide insights into fundamental attention and utility accumulation processes, some uncertainty remains about the extent to which they can be generalized to other settings as we discuss in section 5.4.3.

5.4.1 Collaborative Filtering

Collaborative filtering assumes that consumers with similar past purchases have similar preferences. Therefore, these algorithms recommend items that the focal user has not yet bought but other similar users have. This approach does not require the algorithm to know the content of the recommended items or how users evaluate their features, which makes the algorithm relatively simple to implement for online retailers with a large customer base and assortment. The models we develop in chapter 2 and 3 are also agnostic to the specific features of the brands and predict choice from how consumers attend to the brands from moment-to-moment. To the extent that the results of our essays extend to browsing data, then online retailers who are using collaborative filtering algorithms could benefit from augmenting purchase data with information about which brands the consumer has inspected thus far, the duration, and the pattern.

Amazon has been using collaborative filtering since 2003 and the algorithm that they developed also forms the basis of recommender systems implemented by other companies (e.g. YouTube) (Smith and Linden 2017). A specific aspect of the Amazon recommender system if that after the algorithm identifies similar items from which to make a recommendation, those items that have already been seen by the focal consumer are removed from the set of possible recommendations(Smith and Linden 2017). The underlying assumption seems to be that consumers follow an optimal search process by inspecting one brand at a time and then moving on to a different brand if the current option is not going to be chosen (De los Santos, Hortaçsu, and Wildenbeest 2012; Weitzman 1979). However, given the results of this dissertation, the decision to no longer recommend previously seen items is surprising. Specifically, we find that consumers frequently revisit the brand that they eventually choose even when the information for all of the brands is available at the same time on the screen. This pattern is not specific to eye movements and has also been documented for browsing data (Bronnenberg et al. 2016).

We can only speculate on the effects of this particular aspect of the Amazon recommender system. To the extent that recommendations for yet-unseen items allow the consumer to make a more informed choice after inspecting more diverse alternatives, and consumers appreciate the novelty, this could be a good strategy. However, these recommendations are based on items frequently bought by the same type of consumer and do not include information about the content of the items. Hence, it is not obvious that the recommended items facilitate a more informed choice while it is fair to assume that exploring new items extends decision time. To the extent that the results of our essays extend to browsing data, then online retailers could benefit from predicting (1) the choice likelihood of previously seen items and (2) the remaining time until the consumer is expected to make a purchase. These predictions would then inform the decision of whether an item that has already been seen should be recommended or not.

5.4.2 Content-based Filtering

Content-based filtering uses information about the content of the items to make recommendations and is mostly used by music and movie-streaming platforms. For example, Netflix uses a combination of content-based and collaborative filtering. On the home page, users are recommended movies similar to what they have previously watched on the platform. When users enter a search query, the results are based on the actions of others who have entered similar queries. Ideally, Netflix users are able to find a movie they like within seconds or minutes. Usually these choices are made by inspecting the visual information about the available alternatives and users spend very little time on each movie (1.8 seconds) (Lamkhede and Das 2019).

While the results of the models developed in chapters 2 and 3 are relevant for collaborative filtering recommender systems, the model proposed in chapter 4 includes information about attribute-types across product categories which makes the results of this chapter more relevant for content-based recommender systems. An interesting aspect of choices on Netflix is that users aim not only to choose a good movie, but also that they choose this movie within a short amount of time. Therefore, the company needs to optimize the recommendations offered to its users not only for the utility they derive from watching the movie (i.e. making sure that Netflix users find a good movie to watch), but also for the utility of the search process (i.e. making sure that users don't spend too much time before finding a good movie, but also not too little time as this might lead to an insufficient exploration of available options).

There are several issues that companies such as Amazon and Netflix highlight when discussing recommender systems. First, what recommendations to offer to a new user for whom only limited information about preferences is available (Lamkhede and Das 2019; Smith and Linden 2017). For example, Netflix *jump-starts* their recommendations by asking new users to select a few movie titles that they like. Once users watch content on the Netflix platform, more recent consumption moments receive a larger weight when making future recommendations. Second, how to include data about *when* previous purchases have been made or *when* certain brands have been looked at. This is important in the calibration step. Third, at what time in the future is it best to make a recommendation? We think that the answers to these questions are extremely valuable to companies. At the same time, we argue in favour of a balance between a pure optimization or machine learning approach in which a specific outcome of interest is maximized (usually sales) and a causal approach that aims to uncover fundamental processes that govern how recommendations work. This would allow companies to develop tools that maximize the value they offer to consumers over a longer time horizon as opposed to the usual short-term focus on maximizing sales. This becomes even more important if we consider the recent increased interest in consumer privacy, as we discuss in the next section.

5.4.3 Extensions to Other Types of Display and Product Categories

Mobile devices. Participants in the studies presented in this dissertation have made choices from a computer monitor. Smartphones have a much smaller screen size, which means that consumers need to scroll or click to be able to see more details about products. This leads to at least two differences between the devices that participants in our studies have used and smartphones. First, because scrolling is more effortful and deliberate than eye movements, it should also reflect higher levels of utility. Second, when using a smartphone, previously inspected information is unlikely to still be available on screen. Therefore, memory decay can play a more important role.

While consumers are increasingly using mobile devices, they are less likely to use them for expensive purchases. A study of more than 37K consumers’ clickstream data during one year found that conversion rate increases when consumers switch from a more mobile device to a lesser one, especially when risk (e.g. functional, financial, privacy) is high (De Haan et al. 2018). In the first quarter of 2019, the average online order placed from a mobile phone was 78USD while from a desktop was 119USD[11]. When using a mobile device, consumers search information not only on the shopping website, but on other apps or platforms (e.g. YouTube, Google) as they "are no longer following a linear path from awareness to consideration to purchase"[12].

Websites that do not display attribute by brands matrices to enable product comparison. Even when there are no attribute-by-brands matrices, consumers are presented with multiple items at the same time, displayed either in a list or in a grid display. Therefore, consumers still receive information about multiple brands at the same time on the screen, albeit in a different format. If they want to know more details about a brand, they can click it and are redirected to another page. This would imply that (1) consumers would need to rely more on their memory in order to make brand comparisons, and that (2) as compared to the participants in our studies, they would be more likely to form overall evaluations of a brand earlier in the process, thus more in line with a sequential search approach. A study on search and choice for digital cameras (using browsing and choice data) found that consumers frequently revisit those brands that they eventually choose (Bronnenberg et al. 2016). This suggests that consumers do not rely exclusively on their memory and also do not form overall evaluations of a brand to the extent predicted by sequential search and choice models.

Low involvement product categories. Participants in our studies made choices in six different product categories (Table 5.1). While four of the six categories contain complex brands (e.g. smartphones), the remaining two are in relatively simple categories (lightbulbs and travel mugs). Regardless of the product category, the results show that attention, utilities,

---

[11] https://www.statista.com/statistics/239247/global-online-shopping-order-values-by-device/
[12] https://www.thinkwithgoogle.com/feature/search-intent-marketing-funnel/#/

and brand choice probabilities are closely aligned. This allows our model to predict brand choice early on (after 25-35% of the decision time), irrespective of the duration of the task (self-controlled and relatively longer in Chapter 2 and 3, experimentally manipulated and shorter in Chapter 4). Applications of the aDDM provide additional support for the general nature of the link between attention and choice. These models are usually calibrated on choice between very simple items (e.g. snacks) for which involvement is expected to be lower than for the product categories used in our studies. While decision time is very short (2-3 seconds), the link between attention and choice probability remains (Krajbich et al. 2010).

Choice-based conjoint (CBC) studies. CBC is often used to measure consumer preferences and to segment the market. This method requires participants to make repeated choices within the same category between ‘profiles’ – combinations of attribute levels that are experimentally manipulated. Depending on the number of attributes and attribute levels, the number of choice sets or tasks varies between eight and 20 (Johnson and Orme 1996). Then, exploiting heterogeneity in partworth estimates, segments that group consumers with relatively homogenous preferences are grouped together (e.g. price sensitive consumers, quality oriented, first-time buyers/experimenting). Ideally, the people in these segments are different not only in terms of unobservable preferences, but also other variables that are more easily accessible to marketeers and that they can use to customize their offers, communication strategy, and thus target the right consumer with a matching offer/message. In order for segmentation and profiling to provide valuable insights for companies, preferences elicited in conjoint tasks should be the same as those that consumers have at the point of purchase in a realistic scenario. If that is so, then preferences during the conjoint study are also stable. However, the results of a CBC study that used both choice and eye movement data suggest that early vs. later choice tasks reveal different consumer preferences (Meißner et al. 2016).

5.5 Implications for Consumer Protection and Policy

Consumers use various technologies in their everyday life, such as mobile applications, internet browsers, social networks, and wearables. In doing so, consumers leave a digital footprint - a stream of data that describes their activities. While eye-tracking data is not yet easily available, other types of data that reflect attention and consumer interest are. These types of data (e.g. browsing history, geo-location, purchase history, product ratings, product reviews) are routinely collected online, and in recent years it has become apparent that such practices raise important privacy and ethical concerns. Legislation related to these issues has been adopted within the EU (General Data Protection Regulation) and is also being drafted by other countries outside EU. In this section we discuss: (1) digital footprints as a data source that reflects consumer preferences, (2) technological developments that make is easier to eye track consumers using their own devices (smartphones, laptops), and (3) legislation initiatives that focus on privacy and data protection.

5.5.1 Digital Footprints

Often, consumers explicitly share content with other people (e.g. social networks, forums) or companies (e.g. when making an online purchase, contacting customer service). Data that consumers intentionally submit online is defined as active digital footprint. At the same time, consumers also leave a passive digital footprint: their search history (on Google or through the product catalogue of an online retailer), the news they read online, the location and time when they use a device. Both active and passive digital footprints can reveal consumer attitudes, interests, and preferences (Ghose, Ipeirotis, and Li 2018). This brings numerous opportunities and challenges for companies, consumers, and public policy makers.

Marketing analysts benefit from the variety and volume of data that consumers generate online as this provides very detailed information at individual level. Such data can be used to improve measurement of consumer engagement, preferences, and attitudes. Brand managers can use these insights to gain a better understanding of the consumer and to adapt their communication and branding strategies accordingly. Before companies had access to individual consumers' digital footprints, they used data from panels of households to inform their marketing strategy. Even though these panels contain far fewer data points per household, they could still be used to increase revenues and optimize targeting campaigns (Rossi, McCulloch, and Allenby 1996). Companies' increasing access to detailed information about individual consumers makes it easier to continuously optimize targeting and reach specific consumer segments (e.g. the Facebook ad platform facilitates the delivery of targeted ads based on very specific characteristics of the consumers). Developments in facial recognition technology make it possible to optimize advertising during sport events in real time, based on the composition of the audience (e.g. gender, age) and their engagement with the content. The company that offers this service, Fancam, describes it as "the ultimate crowd-selfie and while engagement is facilitated through cutting edge technology, the real driver is tribalism and our need as individuals to say: 'Look, that's me.'"[13].

Ad effectiveness is influenced not only by the accuracy of targeting consumers, but also by ad transparency and platform trust (Kim, Barasz, and John 2018). This suggests that platforms can derive financial benefits if they build trust with their users as being transparent about targeting criteria on trusted platforms increases advertising effectiveness. However, more research into these topics is needed to determine how privacy concerns, transparency about personalized targeting (ads, recommendations), and the effectiveness of marketing actions are related. In light of their users' growing interest in privacy, companies have

[13] http://www.fancam.com/, accessed 20 September 2019.

implemented different tools that allow users more control over what ads they are exposed to and more information about the ads that they receive. Users (e.g. Facebook, YouTube, Twitter) can access some minimal information about why they are being shown a certain ad, if they choose to click on 'why this ad' buttons displayed next to the ad.

5.5.2 Eye-tracking Solutions and Recent Developments

The studies in this dissertation use eye tracking data collected in two different settings: research areas inside shopping centers (Chapter 2) and in the lab (Chapters 3 and 4). Participants in these studies inspected information presented on a display that looks very similar to a regular PC monitor. In addition to the screen-based eye trackers used in our studies, there are other products such as glasses or VR headsets (Appendix G provides examples from one of the leading companies in this market). While there are several methods to detect eye movements, the general approach is to use (near) infrared light to illuminate the eyes of the user and then a camera to capture how the eyes reflect the light. Many smartphone and laptop models have infrared emitters and cameras that are currently used for facial recognition or iris scanning. For example, iPhone X and later, iPad Pro models with A12X Bionic Chip[14]; laptops that support facial recognition (Windows Hello); smartphones that use iris scans (Samsung Galaxy S8 / Note 8 or later[15]); Dell XPS laptops have two infrared emitters which enable the infrared camera to sense and track motion. Such technological developments make it possible to collect eye tracking data using a user's own device.

Large companies such as Facebook, Google, Unilever have been using eye tracking studies to understand how consumers use their services and react to their advertisements. Both Facebook[16] and Google[17] partnered with Tobii Pro Insight to understand how consumers

---

[14] https://support.apple.com/en-us/HT208108
[15] https://news.samsung.com/global/in-depth-look-keeping-an-eye-on-security-the-iris-scanner-of-the-galaxy-note7
[16] https://www.facebook.com/business/news/insights/mobile-and-tv-between-the-screens
[17] https://www.tobiipro.com/blog/google-eye-tracking-study-tv-youtube-ads/

allocate their attention to advertising while using multiple devices (TV and mobile). Facebook has also been applying for patents that develop eye tracking solutions. For example: a patent on automatic eye tracking calibration that takes place while an individual uses their device (Lopez et al. 2017), and a second one on how to use a device's camera to monitor emotions and display content accordingly (Naveh 2017). When asked about these patents, by the US Senate Committee on Commerce, Science and Transportation, Facebook stated that the company was at that time (June 8, 2018) only exploring how new technology and methods can improve their services and was not building technology to identify people with eye-tracking cameras[18].

5.5.3 Data Protection: Legislation Initiatives

The GDPR defines biometric data as "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural persona, which allow or confirm the unique identification of that natural person" (Article 4(14)) and prohibits processing of biometric data for the purpose of uniquely identifying a natural person (Article 9(1)). Iris texture and facial features fall under the definition of biometric data and could potentially be collected at the same time that eye movements are tracked. While eye movement data would not allow the unique identification of a person, it would make it easier to identify a user within a household. Currently, if members of a household use a shared device without signing in with their own separate account, the digital footprint can be incorrectly assigned to the account that is logged in. Such incorrect assignments could be resolved by verifying that the interpupillary distance of the user matches the account.

Consumers are using facial recognition and fingerprint reading to open their device, although the less intrusive alternative of using a password is still available. It seems that the

[18] https://www.judiciary.senate.gov/meetings/facebook-social-media-privacy-and-the-use-and-abuse-of-data

convenience of logging in by scanning one's face or finger offers enough utility to compensate for the loss of privacy and potential risks of data breaches. Measuring the value of one's data is a difficult task, in part because of the information asymmetry between the companies that collect it and the users that generate it. Therefore, asking users for their willingness to pay (WTP) for using a product (i.e. this could be a device, a feature, an online platform) or willingness to accept (WTA) to forgot it, two measures commonly used to assess welfare (Hanemann 1991) is unlikely to offer a good estimate of the financial value of their data. Even though standard economic theory would predict similar WTP and WTA, the two differ substantially in the context of online platforms such as Facebook, WhatsApp, or LinkedIn. A survey with a US sample of online platform users found that respectively, 33%, 15%, and 23% of them would not be willing to pay for an account (Sunstein 2019). Those participants who were willing to pay for an account offered an average WTP of 26$, 41$, and 33$ (per month) respectively. The reported WTA were more than double: 75$, 101$, and 98$. Another study that measured WTA for Facebook with an incentive compatible design has found an even larger value of 126$ (Corrigan et al. 2018).

Another way of measuring the value of one's data is to base it on the price that advertisers need to pay in order to target a user group. This approach highlights the fact that not all data have the same financial value. The budget of advertisers interested in a user with certain characteristics has a direct impact on the price that Facebook or other advertising platforms can charge for placing an ad in front of that user's eyes. All else equal, the data of users who control household purchases is more valuable (Lambrecht and Tucker 2019). Similarly, the value of data could depend on a large number of factors: income, life events associated with increased spending (e.g. becoming a parent, getting a divorce, job promotions), location of primary residence, health status. Given the sensitive nature of some of these types of data, legislation that limits the collection and use of consumer data has been adopted in the EU (General Fata Protection Regulation[19]) and California (California Consumer Privacy Act[20]). In addition, a legislative initiative (Designing Accounting Safeguards to Help Broaden Oversight and Regulations on Data (DASHBOARD) Act) that focuses specifically on the financial value of user data has recently been introduced by two US senators[21]. If adopted, this bipartisan legislation would require companies to disclose to users an assessment of the value of their data.

5.6 Next Steps Towards a Theory of Rational Attention

The first chapter of this dissertation argues that consumers decide what to focus their attention on, rather than what to be inattentive to. While this is closer to rational *attention*, the models developed in the three empirical essays of this dissertation build on and refer to RIT. We now come back to this idea and propose a theory of rational attention (TRA). TRA specifies that attention, rather than inattention or information acquisition, has an important role for utility accumulation processes that take place during brand choice. We provide a set of propositions (Table 5.3) that capture important aspects of TRA and discuss how these are supported by the results of this dissertation and by previous research. However, there are many questions that remain unanswered as we indicate in the description of each proposition. We argue that investigating these questions are interesting avenues for future research and speculate that this would take us closer to a Theory of Rational Attention.

[19] https://eur-lex.europa.eu/eli/reg/2016/679/oj Regulation (EU) 2016/679 of the European Parliament and of the Council, of 27 April 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
[20] https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375
[21] https://www.warner.senate.gov/public/index.cfm/2019/6/warner-hawley-introduce-bill-to-force-social-media-companies-to-disclose-how-they-are-monetizing-user-data

Table 5.3 Propositions for a theory of rational attention

1. Attention is a coordinating mechanism through which consumers implement their decision goals.

2. Attention is different from other related processes that can also be reflected in eye movements (information acquisition and processing) or lack thereof (inattention).

3. Differences in attention between locations on the display reflect preferences for the items displayed in those locations.

4. Consumers inspect the information on display if they expect that this allows them to reduce uncertainty about their choice.

Proposition 1

Attention is a coordinating mechanism through which consumers implement their decision goals.

An important first step towards a theory of rational attention is specifying what attention is and explaining how it is different from other processes that take place during choice. In this dissertation, *attention* denotes covert attention, “the taking possession by the mind, in clear and vivid form, of one out of [...] several simultaneously possible objects” (James et al. 1890, p. 404). When consumers inspect brands presented on a visual display, their attention is closely related to the eye movements they make during the task (Liechty et al. 2003). While eye movements are not error-free indicators of attention (section 2.2.1), they are reliable reflections of what consumers are focusing on (Pieters and Wedel 2007; Wedel et al. 2008).

Illustrated by previous research on goal control in advertising (Pieters and Wedel 2007) and supported by the results of chapter 4, consumers use eye movements to inspect areas of a visual display that contain information relevant for their task. These results are based on studies in which goals are manipulated through explicit instructions offered to participants. Hence, it is not clear to what extent more subtle attempts at influencing consumers' decision goals would lead to similar effects on attention. For example, when consumers inspect brands in a display their attention can be briefly captured by characteristics of the display (e.g. luminance, color, position) (van der Lans et al. 2008b). Under what conditions can attention initially captured by bottom-up effects can be used to influence the decision goals of the consumer? Then, to the extent that bottom-up effects on attention impact attribute importance or subjective brand values, and knowing that that consumers focus on different areas of a display which depend on their initial decision goals and motivation, how should actions aimed at influencing their goals be adapted?

Proposition 2

> Attention is different from other related processes that can also be reflected in eye movements (information acquisition and processing) or lack thereof (inattention).

Eye-fixations are brief moments when the eyes are relatively still (for about 200-400 msec.) and focused on a specific location in space to acquire information from it. During these moments the consumer can acquire information by reading what is presented on the display (Rayner 1998). Once the consumer acquires some information, deliberation is needed to process it. A distinction is made between information-acquisition and information-processing due to additional cognitive costs that consumers incur when processing information (Payne and Bettman 2004). For example, a consumer who acquires information about the annual energy consumption of a dishwasher uses additional resources (cognitive effort, time) in order to understand if that is an environmentally friendly option or not. During a task, eye movements that consumers make reflect what they focus their attention on (Liechty et al. 2003). In the previous example, a consumer who repeatedly fixates on energy consumption information is more likely to pay more attention to energy efficiency as compare to others who only briefly inspect this attribute. In the absence of eye fixations on a specific area of the display it is reasonable to infer that the consumer did not acquire, process, or attend to the corresponding information (brand and/or attribute). However, the sheer presence of some fixations on a specific area of the display is not a guarantee that the consumer acquired, processed, or attended to the corresponding information. This has several implications. First, eye movements are a necessary but not sufficient condition for information acquisition and processing to occur. Second, information acquisition is a precondition for information processing. Third, consumers' attention to a specific type of information does not guarantee that the information is acquired and/or processed. For example, a consumer whose attention is focused on the environmental impact of their choice is likely to try to acquire relevant information. However, this consumer could fail to find and/or understand this information.

Inattention in rational inattention theory (RIT) is best described by notions of information theory in the communication of messages, specifically Broadbent's (1958) selective filter theory (Masatlioglu et al. 2012). The implication is that decision makers filter out some information in a way similar to how in a noisy environment one needs to make an active choice about what conversation to follow while ignoring other conversations around them. RIT models specify that the decision makers restrict attention to one information acquisition strategy that they use to make a choice (Steiner et al. 2017). If this is the case, then the role of attention is that of "early selection" or a gate that makes sure some information is filtered out. If attention were just a gate or filter that would let information enter the decision process or not, then inattention would be its complement[22]. However, the

[22] In set theory, the complement of a set A refers to elements not in A. Let X be the set containing all the information on display and A be the set containing all the information that a consumer attended to. Then, $IA = X \setminus A$ is the complement of A. IA contains all the information that the consumer did not attend to.

assumption of the eliminative filter is not supported by empirical evidence and "information flowing in unattended channels is not switched off but simply weakened or attenuated" (LaBerge 1995, p. 24).

Some important questions remain. First, are there consumer choice tasks where only attention, inattention, or information acquisition/processing play the main role while the others are less important? We argue that attention is always the most relevant, as it is more informative about what consumers are interested in. However, the empirical results of this dissertation are based on choice tasks in a specific domain: choice between complex brands for which information is displayed at the same time in an attributes-by-brands matrix. Future research could investigate the extent to which the effects we find generalize to other decision types and domains. For example, it would be interesting to examine situations when consumers (1) construct their own choice sets by actively including or excluding brands from a larger, awareness set (Shocker et al. 1991) or (2) make choices best described as a mix of stimulus- and memory-based (Lynch Jr, Marmorstein, and Weigold 1988). In terms of different decision domain, future applications of these models could focus on (1) financial or investment decisions (e.g. what information do investors ignore and what do they pay attention to? Is the performance of investments decisions explained by what investors paid attention to or by what they ignored?), (2) medical decisions, and (3) voting decisions.

While RIT models assume that decision makers attend only to information that reduces their uncertainty, there are situations when decision makers are in fact actively avoiding information that would do so (Gigerenzer and Garcia-Retamero 2017). Such willful ignorance impacts not only brand choice (Ehrich and Irwin 2005), but more importantly medical choices as well. For example, the choice of taking a genetic test to determine the risk of developing cancer (Biesecker et al. 2000) or vaccination choices (Maayan-Metzger, Kedem-Friedrich, and Kuint 2005; Motta, Callaghan, and Sylvester 2018).

Proposition 3

> Differences in attention between locations on the display reflect preferences for the items displayed in those locations.

The three empirical essays in this dissertation find that differences in attention between brands and between attributes are closely linked to the subjective value that consumers attach to those brands and to the importance of the respective attributes. The link between attention and utility has been documented in a variety of other settings (Glaholt et al. 2009; Lindsen et al. 2010; Pieters and Wedel 2007; van der Lans et al. 2008b). We argue that this is a fundamental aspect of attention and therefore expect it to be manifest in other situations. Because the models developed in this dissertation are agnostic to the specific information that is being attended to, we argue that they can be used to also for choice between non-comparable brands (Bettman et al. 1998; Shocker et al. 1991). For example, this could be applied to model gift choices (Wang and van der Lans 2018) as these are often made among items that are in different product categories.

Another area that would be interesting to explore further is related to the impact that changes during the task have on attention and choice. Participants in the studies of this dissertation were presented with all the attribute and brand information at the same time. However, it is not clear how the effects we find in the three empirical essays would change if the moments when participants inspect specific brand and attribute information are not completely under their control, but are also influenced externally. This becomes even more relevant in the context of actions that online retailers could take to influence the choice of a consumer who is in the search and choice process. More specifically, future research could examine: (1) how changes in the amount and type of brand-and-attribute information influences attention and choice, (2) whether changes in information specific to one brand (e.g. price discount for that one brand) impact the evaluation of all the brands in the set, and (3) what processes capture the link between changes in information and utility accumulation.

Proposition 4

Consumers inspect the information on display if this allows them to reduce uncertainty about their choice.

During brand choice, consumers inspect information about the brands in order to reduce uncertainty about their utility match. While this idea is shared by various types of choice theories, what each of these theories assumes by "uncertainty reduction" can differ considerably, with important implications for how preferences are measured and how predictions for brand choice and moment of choice are made (if possible), as discussed in sections 1.3.1 and 5.3.1.

Models developed in the optimal sequential search and choice theoretical framework (Weitzman 1979) assume that consumers already know the brand attributes. Therefore, consumers engage in a search process in order to reduce type III uncertainty about the overall utility of the brand (Kim, Albuquerque, and Bronnenberg 2017). However, it seems rather unlikely that consumers know all the brand attributes before they start inspecting the brands. This explains why models that account for the moment-to-moment reduction in both components of brand uncertainty have found that this leads to more efficient preference measurements (Yang et al. 2015).

Search and choice models (De los Santos et al. 2012; Kim et al. 2017; Weitzman 1979) and models of bounded rationality (Gabaix et al. 2006; Yang et al. 2015) both assume that uncertainty is at the brand level, but not at the choice set level. Specifically, consumers inspect a brand in order to reduce uncertainty about the utility of that focal brand. Once the uncertainty about the utilities of two or more brands is eliminated, the consumer is certain of their utility differences. Sequential sampling models, such as the aDDM, make a different assumption – consumers spend time inspecting the brands because they are unsure about the difference in utilities between the brands (Krajbich et al. 2010; Oud et al. 2016). Therefore, uncertainty is at the level of the choice set.

This shows that there are at least two potential explanations for why consumers acquire information: (1) to make an informed choice and (2) to be confident that the chosen brand is the best from the set of B alternatives. The first explanation implies that consumers reduce uncertainty at the brand level, while the second implies that they reduce uncertainty at the choice set level. Importantly, consumers who have different motivations to reduce uncertainty are expected to inspect the brands differently and to focus on different types of attributes. We argue that it is important for model to capture this accurately, especially if these models aim to make predictions about *when* consumers are going to express their choice.

Then, future research could investigate: (1) if the type of uncertainty that consumers try to reduce by inspecting brand information differs between consumers, choice contexts, type of decisions (e.g. memory-based vs. stimulus based (Lynch Jr et al. 1988), and (2) if and how providers of information (e.g. retailers, policy makers, medical practitioners) should highlight information that decision makers have not inspected or considered enough. If consumers are mostly reducing uncertainty at the choice set level, this implies that from moment-to-moment during choice there are two or more brands that are closely competing. Then, studies that ask for provisional choices or that interrupt the task a random times to ask for choice could offer additional insights into (1) how utility accumulates during choice and (2) how consumers adjust decision thresholds that impact the amount of time spent inspecting the brands.

Appendices

A. Chapter 2: Model estimation

Let $y_{ij} = \left(y_{1ij}^{1}, y_{1ij}^{2}, \ldots, y_{tij}^{g}, \ldots,\ y_{Tij}^{G-1}, y_{Tij}^{G}\right)'$ be a vector of length $(G * T)$. Equation 3 becomes:

(A1) $$y_{ij} = X_{ij}\eta_{ij} + e_j + e_{ij}$$

Where X_{ij} is of dimensions $(G * T)$ by $(K * G)$, η_{ij} is a vector of length $(K * G)$, and e_{ij} and e_{ij} are vectors of length $(G * T)$. We assume the unobserved heterogeneities to be normally distributed: $e_j \sim N(0, \Sigma_0)$ (consumer-level) and $e_{ij} \sim N(0, \Sigma_1)$ (brand-level). The variance covariance matrices at the consumer and brand-level have a similar block diagonal structure. We include here the description at the consumer-level:

(A2) $$\Sigma_0 = \begin{pmatrix} \Sigma_{01} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_{0T} \end{pmatrix}$$

Σ_{0t} is a block of size G by G that captures the residual variance for observed eye movements and the covariances between the different saccade types at time t:

(A3) $$\Sigma_{0t} = \begin{pmatrix} \sigma_{0t}^{1} & 0 & 0 & 0 \\ 0 & \sigma_{0t}^{2} & \sigma_{t}^{23} & \sigma_{0t}^{24} \\ 0 & \sigma_{0t}^{32} & \sigma_{0t}^{3} & \sigma_{0t}^{34} \\ 0 & \sigma_{0t}^{42} & \sigma_{0t}^{43} & \sigma_{0t}^{4} \end{pmatrix}$$

We assume normally distributed heterogeneities: $r_j \sim N(0, \Psi_0)$ (consumer-level) and $r_{ij} \sim N(0, \Psi_1)$ (brand-level). The variance-covariance matrices at the consumer and brand-level have a similar block diagonal structure. Ψ is of dimensions $(G * K)$ by $(G * K)$ and contains one block of size $d = 3$ corresponding to trajectories of brand fixations (g=1), and a second block of size $d = 9$ corresponding to trajectories of eye-saccades (g in 2, 3, 4).

B. Chapter 3: Fixations and brand visits

Let $I_{t,i}^{B}$ indicate if the fixation on brand B at time t for individual i is part of a brand visit:

$$I_{t,i}^{B} = \begin{cases} 1 \; if \; \max\left(f_{t-2,i}^{B} + f_{t-1,i}^{B};\; f_{t-1,i}^{B} + f_{t+1,i}^{B},;\; f_{t+1,i}^{B} + f_{t+2,i}^{B}\right) > 0 \\ 0 \; otherwise \end{cases} \quad \text{(B.1)}$$

Where

$$f_{t,i}^{B} = \begin{cases} 1 \; if \; individual \; i \; has \; its \; t^{th} fixation \; on \; brand \; B \\ 0 \; otherwise \end{cases} \quad \text{(B.2)}$$

For example, for an individual with the sequence of fixations ‘b1 b1 b1 b4 b1 b4 b3’ there are two moments (brand visits):

Brand Visit	Brand 1	Brand 2	Brand 3	Brand 4
1	4	0	0	0
2	4	0	0	2

The results of this paper are based on a dataset of 4487 observations (brand visits) for 214 individuals. This dataset is constructed from the raw eye tracking data collected in the experiment. This raw eye tracking data had 27125 fixations for the 214 individuals in the sample. Per individual, the number of fixations varied between 21 and 321, with an average of 127. Starting from this dataset, several steps were taken to clean the data and to aggregate fixations into brand visits. These steps are: (1) Select fixations on the display area that contains brand information, (2) Select fixations with a typical duration ([-2*STD; 2*STD] interval, after log transformation), (3) Select fixations that are part of a brand visit, and (4) Group fixations into brand visits. After the first two steps, the dataset had 22351 fixations that are on the brand information display area and that have a typical duration. The 17.6% of the fixations in the raw dataset that were discarded were either on an empty area of the monitor, or they were of an unusual length (too short or too long), or both. After the third step, the dataset contains 17797 fixations which are part of a brand visit. The 16.8% fixations in the raw dataset that were discarded in this step were isolated fixations on one of the brands, preceded and followed by fixations on different brands. For example, the fixation on brand 2 in the sequence ‘b1 b1 b2 b1 b1’ is isolated because all other fixations are on a different brand.

C. Chapter 3: Heterogeneity in attention trajectories – model fit comparisons

We test the support in our data for heterogeneity in attention at the participant-level and at the brand-level by comparing the fit of the proposed model for the link between eye movements and attention (eqs. 6-8) against four other competing models based on their ability to predict eye-fixation data ($\hat{y}_i^{T_i}$) out-of-sample (section 3.4.4). A model (IAM1 in Table A.1) that restricts all participant- and brand-specific effects and heterogeneities to zero (η_{g0}, Σ_{10}, η_{0b}, and Σ_{01} all fixed to zero) has an expected log predictive density ($\widehat{elpd}_{IAM1}$) of -19176. After including participant (η_{g0}) and brand (η_{0b}) specific effects the fit ($\widehat{elpd}_{IAM2}$) improves to -14955 (brand and participant heterogeneities Σ_{10} and Σ_{01} remain fixed to zero). We test if both participant (Σ_{10}) and brand (Σ_{01}) heterogeneities are supported by the data by comparing the proposed model against models that include only one of the two. The proposed model, which includes both participant and brand heterogeneities provides the best fit (-1866). The model with only brand-level heterogeneity (IAM4: Σ_{01} estimated and Σ_{10} fixed to zero) has a fit of -1882 while the model that includes only participant heterogeneity (IAM3: Σ_{10} estimated and Σ_{01} fixed to zero) has a fit of -14326. The difference between the proposed model and the model with brand heterogeneity only is 15.55 (*s.e.* = 3.59).

Table A.1 Out-of-sample prediction performance supports brand- and participant-heterogeneity in eye movements

	Overall growth η_{00}	Participant effects η_{g0}	Brand effects η_{0b}	Participant Heterogeneity Σ_{10}	Brand Heterogeneity Σ_{01}	ELPD
IAM1	x					-19176
IAM2	x	x	x			-14955
IAM3	x	x	x	x		-14326
IAM4	x	x	x		x	-1882
Proposed model (eq. 8-10)	x	x	x	x	x	-1866

Note: ELPD stands for expected log predictive density. ELPD values closer to zero indicate better fit.

D. Chapter 4: Stimuli

Figure D.1 Choice task 1 stimuli (light bulbs)

	BRAND 1  Sylvania	BRAND 2  Osram	BRAND 3  Philips	BRAND 4  Megaman
Bulb type	Halogen	LED	Halogen	Energy saver
Energy efficiency class	B	A	B	A
Wattage	28 W	10 W	42 W	11 W
Voltage	220-240 V	220-240 V	220-240 V	220-240 V
Light output (lumens)	345 lm	630 lm	803 lm	630 lm
Equivalent to	40 watts	60 watts	60 watts	52 watts
Colour	Warm white	Warm white	Warm white	Daylight
Average lifetime	2000 hours	25000 hours	2000 hours	15000 hours

Figure D.2 Choice task 2 stimuli (travel mugs)

	BRAND 1  Aladdin	BRAND 2  Monbento	BRAND 3  Grace	BRAND 4  Zuperzazial
Volume	350 ml	500 ml	470 ml	400 ml
Size	8.0 x 7.5 x 20.0	6.4 x 6.4 x 19.4	8.7 x 12.7 x 20.3	9.5 x 9.5 x 14.5
Material	Thermo Plastic	Plastic	Thermo Plastic & Double Glass	Bamboo
Recycled material	Yes	No	No	Yes
Weight	270 g	120 g	186 g	200 g

Figure D.3 Choice task 3 stimuli (TVs)

	BRAND 1  Panasonic	BRAND 2  Philips	BRAND 3  Sony	BRAND 4  Samsung
Energy efficiency class	A Energy	A Energy	A+++ Energy	A++ Energy
On-mode power consumption	57 watts	58 watts	41 watts	48 watts
Electricity consumption	83 kWh/year	85 kWh/year	60 kWh/year	70 kWh/year
Screen size	40 inch	42 inch	42 inch	40 inch
Image quality	Full HD	Full HD	Full HD	Full HD
Image resolution	1920 x 1080 pixels	1920 x 1080 pixels	1920 x 1080 pixels	1920 x 1080 pixels
Image motion rate	200 Hz	500 Hz	300 Hz	400 Hz
Audio power	20 watts RMS	24 watts RMS	20 watts RMS	24 watts RMS
Dimensions with stand	95.7 x 63.5 x 29.4	97.7 x 62.9 x 24.0	97.2 x 64.0 x 26.5	95.7 x 61.9 x 21.7

Figure D.4 Choice task 4 stimuli (Combi fridges)

	BRAND 1  Siemens	BRAND 2  Samsung	BRAND 3  Whirlpool	BRAND 4  Bosch
Energy efficiency class	Energy A+	Energy A+++	Energy A++	Energy A+
Electricity consumption	293 kWh/year	172 kWh/year	204 kWh/year	308 kwh/year
Cooling space (litres)	234	245	225	260
Freezer space (litres)	85	98	113	98
Fast chill option	Yes	No	No	Yes
Compartments (fridge / freezer)	2 / 2	2 / 2	2 / 3	2 / 3
Freezer position	bottom	bottom	bottom	bottom
Dimensions (W x H x D)	59.5 x 187.5 x 64.0	59.5 x 178 x 66.8	60.0 x 185.0 x 65.0	60.0 x 201.0 x 65.0
Weight (kg)	101	79	63	94

E. Chapter 4: Utility specifications for brand choice predictions

The model specification for M0 is:

$$u_{ib} = \alpha_b + \varepsilon_{ib} \tag{D.1}$$

Where α_b is the utility of brand b; ε_{ib} is a type I extreme value distributed utility shock.

The model specification for M1 is:

$$u_{ib} = \alpha_b + \alpha_b^{goal} X_i^{goal} + \varepsilon_{ib} \tag{D.2}$$

Where X_i^{goal} indicated the decision goal of participant i ($X_i^{goal} = 1$ for eco-goal and $X_i^{goal} = 0$ for performance-goal); α_b is the utility of brand b for participants in the performance-goal condition; α_b^{goal} is change in utility of brand b for participants in the eco-goal condition; the ε_{ib} is a type I extreme value distributed utility shock.

The model specification for M2 is:

$$u_{ibt} = \theta_{ibt} + \varepsilon_{ibt} \tag{D.3}$$

Where θ_{ibt} is the subjective value of brand b as reflected by the share of attention on brand b ($\bar{\theta}_{ibt}$) relative to the expected share if attention is equally distributed between the brands (25%); ε_{ibt} is a type I extreme value distributed utility shock.

$$y_{ibt} = \bar{\theta}_{ibt} + \xi_{ibt} \tag{D.4}$$

Where y_{ibt} is the observed share of eye-fixations for participant i on brand b at time t as percent of the total number of eye-fixations for participant i at time t; $\bar{\theta}_{ibt}$ is the attention share of participant i on brand b at time t; $\bar{\theta}_{ibt} \sim N(\bar{\theta}_{0bt}, \sigma_{bt}^2)$; $\bar{\theta}_{0bt}$ is the overall attention share for brand b at time t; and σ_{bt}^2 is the participant-level variance in attention shares for brand b at time t.

The model specification for M3 is:

$$u_{ibt} = \alpha_b + \sum_{j=1}^{J} \beta_j \theta_{ibjt} + \varepsilon_{ibt} \tag{D.5}$$

Where α_b is the utility intercept of brand b; θ_{ibjt} is the subjective value of brand b on attribute j at time t; θ_{ibjt} is reflected by the share of attention on brand b and attribute j relative to the fair share of 25%; ε_{ib} is a type I extreme value distributed utility shock.

$$y_{ibjt} = \bar{\theta}_{ibjt} + \xi_{ibjt} \tag{D.6}$$

Where y_{ibjt} is the share of fixations on brand b relative to the total number of fixations on the attribute j for consumer i at time t; $\bar{\theta}_{ibjt}$ is the share of attention on brand b on attribute j for consumer i at time t; $\bar{\theta}_{ibjt} \sim N(\bar{\theta}_{0bjt}, \sigma^2_{bjt})$; $\bar{\theta}_{0bjt}$ is the overall attention share on brand b and attribute j at time t; σ^2_{bjt} is the participant-level variance in attention shares; ξ_{ibjt} unobserved heterogeneity.

The model specification for M4 is:

$$u_{ibt} = \alpha_b + \beta\theta_{ibt} + \varepsilon_{ibt} \tag{D.7}$$

Where α_b is the utility intercept of brand b; θ_{ibt} is the subjective value of brand b as reflected by the share of attention on brand b ($\bar{\theta}_{ibt}$) relative to the expected share if attention is equally distributed between the brands (25%); ε_{ibt} is a type I extreme value distributed utility shock.

$$y_{ibt} = \bar{\theta}_{ibt} + \xi_{ibt} \tag{D.8}$$

Where y_{ibt} is the observed share of eye-fixations for participant i on brand b at time t as percent of the total number of eye-fixations for participant i at time t; $\bar{\theta}_{ibt}$ is the attention share of participant i on brand b at time t; $\bar{\theta}_{ibt} \sim N(\bar{\theta}_{0bt}, \sigma^2_{bt})$; $\bar{\theta}_{0bt}$ is the overall attention share for brand b at time t; and σ^2_{bt} is the participant-level variance in attention shares for brand b at time t.

The model specification for M5 is already provided in Chapter 4, equations 1-6.

F. Chapter 4: Estimation results for choice tasks 1, 2, and 4

Table F.1 Estimation results – attribute importance (choice task 1, lightbulbs)

Choice task 1	Period 1 (γ_{kj1})			Period 2 (γ_{kj2})			Period 3 (γ_{kj3})		
(light bulbs)	*M*	*SD*	p-value	*M*	*SD*	p-value	*M*	*SD*	p-value
Performance-goal ($k = 0$, eq. 7-8)									
Eco-attr.	.04	.02	.02	-.10	.03	<.001	.03	.03	.14
Performance-attr.	.01	.02	.17	.06	.03	.01	-.04	.03	.10
Other attr.	-.05	.02	.004	.04	.03	.06	.02	.03	.18
Eco-goal ($k = 1$, eq. 7-8)									
Eco-attr.	.08	.03	.003	-.01	.04	.58	-.01	.04	.65
Performance-attr.	-.11	.03	<.001	.03	.04	.17	-.01	.04	.64
Other attr.	.03	.03	.12	-.02	.04	.53	.02	.04	.21
High time-pressure ($k = 2$, eq. 7-8)									
Eco-attr.	.07	.03	.004	-.18	.04	<.001	.16	.04	<.001
Performance-attr.	-.14	.03	<.001	.24	.04	<.001	-.16	.04	<.001
Other attr.	.06	.03	.01	-.06	.04	.13	.00	.04	.25
Eco-goal x High time-pressure ($k = 3$, eq. 7-8)									
Eco-attr.	.01	.04	.24	.07	.05	.08	-.10	.05	.06
Performance-attr.	.06	.04	.048	-.16	.05	.003	.14	.05	.004
Other attr.	-.07	.04	.07	.09	.05	.04	-.04	.05	.39

Note: attr = attributes. Estimates in the "Period 2" column indicate changes in attribute attention shares relative to the "Period 1" column. Estimates in the "Period 3" column indicate changes in attribute attention shares relative to the "Period 1" and "Period 2" columns. Shares within a condition and time interval sum to 0. The values are as compared to the "fair share" based on the number of attributes of each type as compared to the total number of attributes. See table 4.1.

Table F.2 Estimation results – attribute importance (choice task 2, travel mugs)

Choice task 2 (travel mugs)	Period 1 (γ_{kj1})			Period 2 (γ_{kj2})			Period 3 (γ_{kj3})		
	M	*SD*	p-value	*M*	*SD*	p-value	*M*	*SD*	p-value
Performance-goal ($k = 0$, eq. 7-8)									
Eco-attr.	-.07	.02	<.001	.06	.03	.01	-.04	.03	.09
Performance-attr.	.04	.02	.01	.04	.03	.05	-.06	.03	.01
Other attr.	.02	.02	.08	-.11	.03	<.001	.11	.03	<.001
Eco-goal ($k = 1$, eq. 7-8)									
Eco-attr.	.00	.02	.25	.03	.03	.16	-.01	.04	.64
Performance-attr.	.03	.03	.15	-.06	.04	.09	.03	.03	.16
Other attr.	-.03	.02	.24	.03	.03	.14	-.02	.03	.50
High time-pressure ($k = 2$, eq. 7-8)									
Eco-attr.	.00	.03	.25	.02	.04	.20	-.03	.04	.46
Performance-attr.	.01	.03	.24	.05	.04	.10	-.01	.04	.68
Other attr.	-.01	.03	.58	-.07	.04	.06	.03	.04	.16
Eco-goal x High time-pressure ($k = 3$, eq. 7-8)									
Eco-attr.	-.04	.04	.22	.04	.05	.18	.03	.05	.20
Performance-attr.	-.03	.04	.42	.04	.05	.16	-.04	.05	.38
Other attr.	.07	.04	.02	-.08	.05	.11	.01	.05	.24

Note: attr = attributes. Estimates in the “Period 2” column indicate changes in attribute attention shares relative to the “Period 1” column. Estimates in the “Period 3” column indicate changes in attribute attention shares relative to the “Period 1” and “Period 2” columns. Shares within a condition and time interval sum to 0. The values are as compared to the “fair share” based on the number of attributes of each type as compared to the total number of attributes. See table 4.1.

Table F.3 Estimation results – attribute importance (choice task 4, fridges)

Choice task 4 (fridges)	Period 1 (γ_{kj1})			Period 2 (γ_{kj2})			Period 3 (γ_{kj3})		
	M	*SD*	p-value	*M*	*SD*	p-value	*M*	*SD*	p-value
Performance-goal ($k = 0$, eq. 7-8)									
Eco-attr.	.02	.02	.14	-.08	.03	<.01	.05	.03	.03
Performance-attr.	.06	.02	<.001	.00	.03	.70	-.03	.03	.25
Other attr.	-.07	.02	<.001	.08	.03	<.001	-.02	.03	.41
Eco-goal ($k = 1$, eq. 7-8)									
Eco-attr.	.17	.02	<.001	-.16	.04	<.001	.08	.03	.01
Performance-attr.	-.11	.02	<.001	.10	.04	.002	-.04	.03	.23
Other attr.	-.06	.02	.01	.07	.03	.03	-.04	.03	.25
High time-pressure ($k = 2$, eq. 7-8)									
Eco-attr.	.01	.03	.24	-.02	.04	.50	.01	.04	.22
Performance-attr.	-.02	.03	.45	.01	.04	.24	.01	.04	.24
Other attr.	-.01	.03	.23	.01	.04	.24	-.03	.04	.40
Eco-goal x High time-pressure ($k = 3$, eq. 7-8)									
Eco-attr.	.13	.04	<.001	-.08	.05	.14	-.05	.05	.27
Performance-attr.	-.01	.04	.67	.00	.05	.73	.01	.05	.24
Other attr.	-.12	.04	<.001	.08	.05	.05	.04	.05	.17

Note: attr = attributes. Estimates in the “Period 2” column indicate changes in attribute attention shares relative to the “Period 1” column. Estimates in the “Period 3” column indicate changes in attribute attention shares relative to the “Period 1” and “Period 2” columns. Shares within a condition and time interval sum to 0. The values are as compared to the “fair share” based on the number of attributes of each type as compared to the total number of attributes. See table 4.1.

Table F.4 Estimation results – subjective values of brand and attribute combinations (choice task 1, light bulbs)

Choice task 1 (light bulbs)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Performance-goal ($k = 0$, eq. 9-10)										
Eco-attributes	A	-.05	.02	.03	-.09	.03	.01	.02	.03	.20
	B (Eco)	.04	.02	.048	-.06	.03	.10	.10	.03	<.001
	C (Perf)	.04	.03	.07	.17	.03	<.001	-.12	.03	<.001
	D (Eco)	-.02	.02	.29	-.06	.04	.08	.04	.03	.08
Performance-attributes	A	-.11	.02	<.001	.01	.03	.25	-.01	.03	.56
	B (Eco)	-.03	.02	.18	.01	.03	.22	.04	.03	.11
	C (Perf)	-.01	.02	.59	.16	.03	<.001	-.07	.03	.04
	D (Eco)	-.09	.02	.00	.03	.03	.17	.00	.03	.25
Other-attributes	A	-.07	.02	.003	-.05	.03	.17	-.03	.03	.40
	B (Eco)	.06	.03	.01	-.08	.03	.03	.10	.03	.002
	C (Perf)	.02	.02	.15	.19	.04	<.001	-.09	.03	.01
	D (Eco)	-.02	.02	.38	-.05	.03	.14	.02	.03	.21
Eco-goal ($k = 1$, eq. 9-10)										
Eco-attributes	A	-.03	.03	.29	.03	.05	.19	.02	.05	.23
	B (Eco)	.03	.03	.14	.09	.05	.02	-.03	.05	.39
	C (Perf)	-.07	.03	.03	-.21	.05	.00	.10	.05	.01
	D (Eco)	.07	.03	.02	.11	.05	.01	-.10	.05	.02
Performance-attributes	A	-.06	.03	.05	.04	.05	.17	.03	.05	.19
	B (Eco)	.12	.03	<.001	.04	.05	.16	.00	.05	.72
	C (Perf)	-.15	.03	<.001	-.11	.05	.02	.06	.05	.09
	D (Eco)	.07	.03	.01	.02	.05	.21	-.04	.05	.32
Other-attributes	A	.01	.03	.24	-.02	.05	.51	.02	.05	.22
	B (Eco)	.06	.03	.04	.13	.05	.00	-.09	.05	.045
	C (Perf)	-.14	.03	<.001	-.17	.05	.00	.08	.05	.04
	D (Eco)	.08	.03	.01	.05	.04	.13	-.01	.05	.66
High time-pressure ($k = 2$, eq. 9-10)										
Eco-attributes	A	.02	.04	.22	-.02	.05	.54	-.01	.05	.63
	B (Eco)	-.07	.04	.04	.04	.05	.15	-.01	.05	.70
	C (Perf)	-.01	.04	.63	-.14	.05	.00	.20	.05	<.001
	D (Eco)	.04	.04	.13	-.01	.05	.65	-.03	.05	.46
Performance-attributes	A	-.09	.03	.01	.06	.05	.08	-.01	.05	.66
	B (Eco)	-.10	.03	.003	.12	.05	.00	-.12	.05	.02
	C (Perf)	-.12	.03	<.001	.05	.05	.14	-.01	.05	.69
	D (Eco)	-.10	.04	.003	.13	.05	.00	-.09	.05	.09
Other-attributes	A	.04	.04	.12	-.04	.05	.37	-.01	.05	.64
	B (Eco)	-.09	.04	.02	.12	.05	.01	-.07	.05	.16
	C (Perf)	.02	.04	.20	-.12	.05	.02	.12	.05	.003
	D (Eco)	-.02	.03	.54	.02	.05	.22	.00	.05	.76

(table continues on next page)

(table continues from previous page)

Choice task 1 (light bulbs)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Eco-goal x High time-pressure ($k = 3$, eq. 9-10)										
Eco-attributes	A	.00	.05	.25	-.03	.07	.51	.02	.07	.23
	B (Eco)	.04	.05	.17	.03	.07	.22	-.04	.07	.47
	C (Perf)	.03	.05	.18	.08	.07	.10	-.14	.07	.03
	D (Eco)	-.04	.05	.34	.03	.07	.21	.02	.07	.24
Performance-attributes	A	.05	.05	.14	-.09	.07	.16	.01	.07	.25
	B (Eco)	-.07	.05	.12	.02	.07	.24	-.02	.07	.61
	C (Perf)	.05	.05	.13	-.01	.07	.72	.03	.07	.21
	D (Eco)	.04	.05	.16	-.04	.07	.49	.01	.07	.25
Other-attributes	A	-.02	.05	.57	.00	.07	.74	.02	.07	.23
	B (Eco)	.00	.05	.25	-.01	.07	.68	.05	.07	.16
	C (Perf)	.00	.05	.71	.07	.07	.12	-.09	.07	.18
	D (Eco)	.04	.05	.18	-.06	.07	.36	-.01	.07	.69

Note: Shares should sum to 0 within a condition, attribute type, and time interval. When shares do not sum to zero it is because some participants do not have fixations on any of the brands for the respective attribute-type within the time interval. Estimates in the "Period 2" column indicate changes in brand-and-attribute attention shares relative to the "Period 1" column. Estimates in the "Period 3" column indicate changes in brand-and-attribute attention shares relative to the "Period 1" and "Period 2" columns. The values are as compared to the "fair share" of .25.

Table F.5 Estimation results – subjective values of brand and attribute combinations (choice task 2, travel mug)

Choice task 2 (travel mug)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Performance-goal ($k = 0$, eq. 9-10)										
Eco-attributes	A (Perf)	-.08	.03	.003	.00	.04	.71	.01	.04	.23
	B	-.09	.03	<.001	.06	.04	.05	-.07	.04	.06
	C (Perf)	-.04	.03	.16	.10	.04	.01	.02	.04	.21
	D (Eco)	-.12	.03	<.001	.06	.04	.04	-.01	.04	.57
Performance-attributes	A (Perf)	-.01	.03	.59	-.02	.04	.56	-.02	.04	.47
	B	-.08	.03	.004	.03	.04	.17	-.02	.04	.46
	C (Perf)	.11	.03	<.001	.04	.04	.13	.01	.04	.25
	D (Eco)	-.09	.03	<.001	.00	.04	.25	-.01	.04	.57
Other-attributes	A (Perf)	.00	.03	.70	-.01	.04	.65	.04	.04	.12
	B	.00	.03	.69	-.01	.04	.67	-.02	.04	.44
	C (Perf)	.04	.03	.06	.05	.04	.07	.00	.04	.25
	D (Eco)	-.03	.03	.21	-.04	.04	.25	-.02	.04	.54

(table continues on next page)

(table continues from previous page)

Choice task 2 (travel mug)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Eco-goal ($k = 1$, eq. 9-10)										
Eco-attributes	A (Perf)	-.05	.03	.16	.16	.05	<.001	-.09	.05	.07
	B	-.02	.03	.44	.00	.05	.75	.06	.05	.09
	C (Perf)	.01	.03	.24	-.13	.05	.02	-.02	.05	.53
	D (Eco)	.11	.04	.00	-.02	.05	.57	.06	.05	.08
Performance-attributes	A (Perf)	.01	.03	.24	.13	.05	.01	-.06	.05	.22
	B	.01	.04	.25	-.07	.05	.13	.04	.05	.15
	C (Perf)	-.06	.03	.10	-.13	.05	.01	.02	.05	.22
	D (Eco)	.04	.03	.09	.08	.05	.05	.04	.05	.17
Other-attributes	A (Perf)	.00	.03	.25	.09	.05	.04	-.11	.05	.02
	B	-.01	.04	.63	-.06	.05	.20	.02	.05	.23
	C (Perf)	-.09	.03	.01	-.09	.05	.07	-.01	.05	.67
	D (Eco)	.10	.0	.00	.07	.05	.07	.11	.05	.02
High time-pressure ($k = 2$, eq. 9-10)										
Eco-attributes	A (Perf)	-.04	.04	.22	.03	.05	.19	-.01	.05	.69
	B	-.08	.04	.045	.02	.05	.22	.04	.05	.19
	C (Perf)	-.07	.04	.06	.00	.05	.25	.05	.05	.13
	D (Eco)	.01	.04	.23	-.04	.05	.39	-.07	.05	.14
Performance-attributes	A (Perf)	-.02	.04	.51	.07	.05	.08	-.01	.05	.70
	B	-.02	.04	.54	-.01	.05	.68	-.01	.05	.68
	C (Perf)	-.14	.04	<.001	.08	.05	.07	.03	.05	.20
	D (Eco)	.02	.04	.21	-.03	.05	.53	-.04	.05	.38
Other-attributes	A (Perf)	.02	.04	.21	-.04	.05	.41	.00	.05	.25
	B	.00	.04	.69	-.02	.06	.61	.04	.05	.19
	C (Perf)	-.04	.04	.21	.05	.05	.15	.01	.05	.24
	D (Eco)	.03	.04	.16	-.03	.05	.47	-.01	.05	.66
Eco-goal x High time-pressure ($k = 3$, eq. 9-10)										
Eco-attributes	A (Perf)	-.05	.05	.29	-.04	.07	.46	.10	.07	.07
	B	-.01	.05	.71	.03	.07	.23	-.06	.07	.36
	C (Perf)	-.05	.05	.30	.13	.07	.03	-.11	.07	.15
	D (Eco)	-.11	.05	.03	.07	.07	.14	.07	.07	.14
Performance-attributes	A (Perf)	-.06	.05	.19	-.07	.07	.30	.07	.07	.14
	B	-.03	.05	.49	.04	.07	.21	-.03	.07	.58
	C (Perf)	.07	.05	.08	.03	.07	.22	-.10	.07	.15
	D (Eco)	-.11	.05	.03	.13	.07	.04	.05	.07	.17
Other-attributes	A (Perf)	-.03	.05	.46	.02	.07	.24	.08	.07	.13
	B	.02	.05	.23	-.03	.07	.56	-.05	.07	.47
	C (Perf)	.07	.05	.08	-.07	.07	.27	-.02	.07	.62
	D (Eco)	-.06	.05	.20	.10	.07	.07	-.02	.07	.63

Note: Shares should sum to 0 within a condition, attribute type, and time interval. When shares do not sum to zero it is because some participants do not have fixations on any of the brands for the respective attribute-type within the time interval. Estimates in the “Period 2” column indicate changes in brand-and-attribute attention shares relative to the “Period 1” column. Estimates in the “Period 3” column indicate changes in brand-and-attribute attention shares relative to the “Period 1” and “Period 2” columns. The values are as compared to the “fair share” of .25.

Table F.6 Estimation results – subjective values of brand and attribute combinations (choice task 4, fridge)

Choice task 4 (Fridge)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	SD	p-value	M	SD	p-value	M	SD	p-value
Performance-goal ($k = 0$, eq. 9-10)										
Eco-attributes	A (Perf)	.00	.03	.71	-.02	.04	.46	.01	.04	.23
	B (Eco)	.02	.03	.16	-.15	.03	<.001	.06	.03	.04
	C	-.05	.03	.05	-.08	.04	.03	.00	.03	.72
	D (Perf)	.02	.02	.18	.10	.03	.002	-.01	.03	.64
Performance-attributes	A (Perf)	-.06	.02	.01	.10	.04	.001	-.06	.04	.12
	B (Eco)	-.06	.02	.01	-.03	.04	.39	.00	.04	.25
	C	-.10	.03	.002	-.02	.04	.49	.03	.04	.14
	D (Perf)	-.07	.02	.01	.09	.03	.01	.03	.04	.14
Other-attributes	A (Perf)	.03	.02	.12	.02	.04	.19	-.02	.03	.51
	B (Eco)	.01	.03	.20	-.09	.04	.02	.00	.04	.73
	C	-.03	.02	.22	-.06	.04	.12	-.01	.04	.59
	D (Perf)	-.01	.03	.58	.12	.04	<.001	.03	.04	.16
Eco-goal ($k = 1$, eq. 9-10)										
Eco-attributes	A (Perf)	-.11	.03	<.001	.01	.05	.24	-.02	.05	.56
	B (Eco)	.15	.03	<.001	.11	.05	.01	.07	.05	.07
	C	.08	.03	.01	-.01	.05	.64	-.01	.05	.63
	D (Perf)	-.11	.03	.002	-.17	.05	<.001	.01	.05	.25
Performance-attributes	A (Perf)	-.15	.03	<.001	.00	.05	.25	.00	.05	.25
	B (Eco)	.10	.03	<.001	.12	.05	.01	-.12	.05	.01
	C	-.01	.03	.62	.07	.05	.08	-.11	.05	.03
	D (Perf)	-.16	.03	<.001	-.01	.05	.63	-.06	.05	.18
Other-attributes	A (Perf)	-.11	.03	<.001	-.02	.05	.54	-.01	.05	.66
	B (Eco)	.21	.03	<.001	.07	.05	.048	.04	.05	.15
	C	.01	.03	.25	.07	.05	.07	.01	.05	.24
	D (Perf)	-.10	.03	.002	-.12	.05	.01	-.04	.05	.35
High time-pressure ($k = 2$, eq. 9-10)										
Eco-attributes	A (Perf)	-.02	.04	.50	-.08	.05	.13	.08	.05	.05
	B (Eco)	-.03	.04	.42	.01	.05	.24	-.04	.05	.43
	C	.04	.04	.14	-.03	.05	.50	-.02	.05	.56
	D (Perf)	-.05	.03	.18	-.12	.05	.02	.14	.05	.003
Performance-attributes	A (Perf)	-.06	.04	.06	-.05	.05	.33	.00	.05	.74
	B (Eco)	-.07	.04	.049	.07	.05	.08	.00	.05	.25
	C	-.06	.04	.11	.05	.05	.14	-.06	.05	.26
	D (Perf)	-.03	.04	.32	-.10	.05	.049	.11	.05	.01
Other-attributes	A (Perf)	.04	.04	.13	-.06	.05	.25	.01	.05	.24
	B (Eco)	-.01	.04	.65	.07	.05	.08	-.06	.05	.23
	C	-.05	.04	.12	.05	.05	.14	.00	.05	.25
	D (Perf)	.01	.04	.24	-.06	.05	.23	.06	.05	.12

(table continues on next page)

(table continues from previous page)

Choice task 4 (Fridge)	Brand	Period 1 (γ_{kbj1})			Period 2 (γ_{kbj2})			Period 3 (γ_{kbj3})		
		M	*SD*	*p*-value	*M*	*SD*	*p*-value	*M*	*SD*	*p*-value
Eco-goal x High time-pressure ($k = 3$, eq. 9-10)										
Eco-attributes	A (Perf)	.06	.05	.11	.02	.07	.24	-.06	.07	.32
	B (Eco)	-.02	.05	.59	.02	.07	.24	-.10	.07	.15
	C	-.06	.05	.23	.04	.07	.20	.03	.07	.23
	D (Perf)	.06	.05	.09	.11	.07	.04	-.14	.07	.05
Performance-attributes	A (Perf)	.05	.05	.15	-.02	.07	.63	.04	.07	.20
	B (Eco)	-.14	.05	<.001	.13	.07	.03	.03	.07	.23
	C	-.02	.05	.54	-.03	.07	.52	.06	.07	.17
	D (Perf)	.04	.05	.17	.07	.07	.12	-.10	.07	.14
Other-attributes	A (Perf)	-.07	.05	.15	.05	.07	.18	.03	.07	.22
	B (Eco)	.03	.05	.19	-.02	.07	.66	-.01	.07	.68
	C	.05	.05	.14	-.06	.07	.34	.00	.07	.25
	D (Perf)	.01	.05	.25	.02	.07	.23	-.03	.07	.57

Note: Shares should sum to 0 within a condition, attribute type, and time interval. When shares do not sum to zero it is because some participants do not have fixations on any of the brands for the respective attribute-type within the time interval. Estimates in the "Period 2" column indicate changes in brand-and-attribute attention shares relative to the "Period 1" column. Estimates in the "Period 3" column indicate changes in brand-and-attribute attention shares relative to the "Period 1" and "Period 2" columns. The values are as compared to the "fair share" of .25.

G. Chapter 5: Eye tracking hardware

Figure G.1 Tobii Pro TX300

Source: https://www.tobiipro.com/product-listing/tobii-pro-tx300/

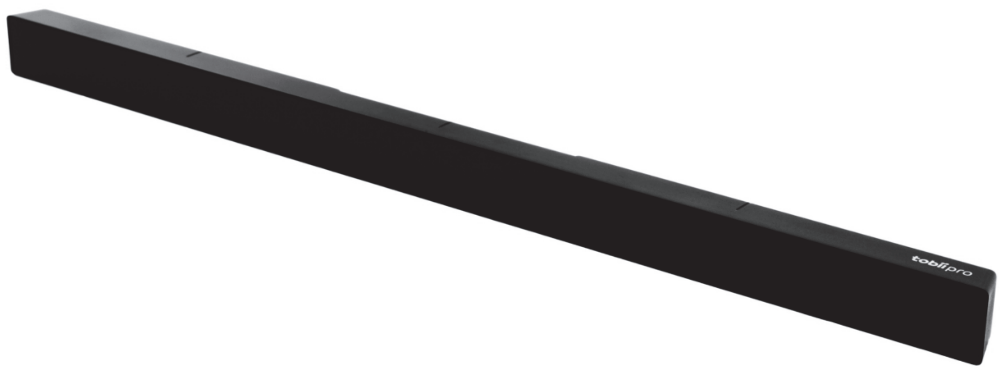
Figure G.2 Tobii Pro Glasses 2

Source: https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/

Figure G.3 Tobii Pro VR Analytics

Source: https://www.tobiipro.com/product-listing/vr-analytics/

Figure G.4 Tobii Pro X3-120



Source: https://www.tobiipro.com/product-listing/tobii-pro-x3-120/

References

Adomavicius, G. and A. Tuzhilin (2005), "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge & Data Engineering*, 17 (6), 734-49.

Al-Moteri, M. O., M. Symmons, V. Plummer, and S. Cooper (2017), "Eye Tracking to Investigate Cue Processing in Medical Decision-Making: A Scoping Review," *Computers in Human Behavior*, 66 (1), 52-66.

Arieli, A., Y. Ben-Ami, and A. Rubinstein (2011), "Tracking Decision Makers under Uncertainty," *American Economic Journal: Microeconomics*, 3 (4), 68-76.

Ashby, N. J., M. Jekel, S. Dickert, and A. Glöckner (2016), "Finding the Right Fit: A Comparison of Process Assumptions Underlying Popular Drift-Diffusion Models," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42 (12), 1982-93.

Atalay, A. S., H. O. Bodur, and D. Rasolofoarison (2012), "Shining in the Center: Central Gaze Cascade Effect on Product Choice," *Journal of Consumer Research*, 39 (4), 848-66.

Berkowitsch, N. A. J., B. Scheibehenne, and J. Rieskamp (2014), "Rigorously Testing Multialternative Decision Field Theory against Random Utility Models," *Journal of Experimental Psychology: General*, 143 (3), 1331-48.

Bettman, J. R., M. F. Luce, and J. W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187-217.

Biesecker, B. B., N. Ishibe, D. W. Hadley, T. R. Giambarresi, R. G. Kase, C. Lerman, and J. P. Struewing (2000), "Psychosocial Factors Predicting BRCA1/BRCA2 Testing Decisions in Members of Hereditary Breast and Ovarian Cancer Families," *American Journal of Medical Genetics*, 93 (4), 257-63.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013), "Salience and Consumer Choice," *Journal of Political Economy*, 121 (5), 803-43.

--- (2016), "Competition for Attention," *The Review of Economic Studies*, 83 (2), 481-513.

Bottomley, P. A., J. R. Doyle, and R. H. Green (2000), "Testing the Reliability of Weight Elicitation Methods: Direct Rating Versus Point Allocation," *Journal of Marketing Research*, 37 (4), 508-13.

Branco, F., M. Sun, and J. M. Villas-Boas (2016), "Too Much Information? Information Provision and Search Costs," *Marketing Science*, 35 (4), 605-18.

Bronnenberg, B. J. and J.-P. Dubé (2017), "The Formation of Consumer Brand Preferences," *Annual Review of Economics*, 9 (1), 353-82.

Bronnenberg, B. J., J. B. Kim, and C. F. Mela (2016), "Zooming in on Choice: How Do Consumers Search for Cameras Online?," *Marketing Science*, 35 (5), 693-712.

Busemeyer, J. R. (1985), "Decision Making under Uncertainty: A Comparison of Simple Scalability, Fixed-Sample, and Sequential-Sampling Models," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (3), 538-64.

Busemeyer, J. R. and J. G. Johnson (2004), "Computational Models of Decision Making," in *Blackwell Handbook of Judgment and Decision Making*, ed. Derek J. Koehler and Nigel Harvey, 133-54.

Busemeyer, J. R. and J. T. Townsend (1993), "Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment," *Psychological Review*, 100 (3), 432-59.

Caplin, A. (2016), "Measuring and Modeling Attention," *Annual Review of Economics*, 8 (1), 379-403.

Caplin, A. and M. Dean (2015), "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105 (7), 2183-203.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell (2017), "Stan: A Probabilistic Programming Language," *Journal of Statistical Software*, 76 (1), 1-32.

Cavanagh, J. F., T. V. Wiecki, A. Kochar, and M. J. Frank (2014), "Eye Tracking and Pupillometry Are Indicators of Dissociable Latent Decision Processes," *Journal of Experimental Psychology: General*, 143 (4), 1476-88.

Chandon, P., J. W. Hutchinson, E. T. Bradlow, and S. H. Young (2009), "Does in-Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase," *Journal of Marketing*, 73 (6), 1-17.

Ching, A. T., T. Erdem, and M. P. Keane (2013), "Learning Models: An Assessment of Progress, Challenges, and New Developments," *Marketing Science*, 32 (6), 913-38.

Corrigan, J. R., S. Alhabash, M. Rousu, and S. B. Cash (2018), "How Much Is Social Media Worth? Estimating the Value of Facebook by Paying Users to Stop Using It," *PLoS One*, 13 (12), 1-11.

Danaher, P. J. and H. J. van Heerde (2018), "Delusion in Attribution: Caveats in Using Attribution for Multimedia Budget Allocation," *Journal of Marketing Research*, 55 (5), 667-85.

De Haan, E., P. Kannan, P. C. Verhoef, and T. Wiesel (2018), "Device Switching in Online Purchasing: Examining the Strategic Contingencies," *Journal of Marketing*, 82 (5), 1-19.

De los Santos, B., A. Hortaçsu, and M. R. Wildenbeest (2012), "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102 (6), 2955-80.

Dean, M. and N. L. Neligh (2019), "Experimental Tests of Rational Inattention," *working paper*, Available at: http://www.columbia.edu/~md3405/Working_Paper_21.pdf.

Diederich, A. (2003), "MDFT Account of Decision Making under Time Pressure," *Psychonomic Bulletin & Review*, 10 (1), 157-66.

Dror, I. E., B. Basola, and J. R. Busemeyer (1999), "Decision Making under Time Pressure: An Independent Test of Sequential Sampling Models," *Memory & Cognition*, 27 (4), 713-25.

Dubé, J.-P., G. J. Hitsch, P. E. Rossi, and M. A. Vitorino (2008), "Category Pricing with State-Dependent Utility," *Marketing Science*, 27 (3), 417-29.

Duffy, S. A., R. K. Morris, and K. Rayner (1988), "Lexical Ambiguity and Fixation Times in Reading," *Journal of Memory and Language*, 27 (4), 429-46.

Ehrich, K. R. and J. R. Irwin (2005), "Willful Ignorance in the Request for Product Attribute Information," *Journal of Marketing Research*, 42 (3), 266-77.

Erdem, T., M. P. Keane, and B. Sun (2008), "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science*, 27 (6), 1111-25.

Fiedler, S. and A. Glöckner (2012), "The Dynamics of Decision Making in Risky Choice: An Eye-Tracking Analysis," *Frontiers in Psychology*, 3 (335), 1-18.

Forstmann, B. U., R. Ratcliff, and E. J. Wagenmakers (2016), "Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions," *Annual Review of Psychology*, 67 (1), 641-66.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg (2006), "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *The American Economic Review*, 96 (4), 1043-68.

Gelfand, A. E. and D. K. Dey (1994), "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society. Series B (Methodological)*, 56 (3), 501-14.

Gelman, A. and D. B. Rubin (1992), "Inference from Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7 (4), 457-72.

Gelman, A., H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013), *Bayesian Data Analysis*: Chapman and Hall/CRC.

Ghose, A., P. G. Ipeirotis, and B. Li (2018), "Modeling Consumer Footprints on Search Engines: An Interplay with Social Media," *Management Science*, 65 (3), 1363-85.

Gidlöf, K., A. Wallin, R. Dewhurst, and K. Holmqvist (2013), "Using Eye Tracking to Trace a Cognitive Process: Gaze Behaviour During Decision Making in a Natural Environment," *Journal of Eye Movement Research*, 6 (1), 1-14.

Gigerenzer, G. and R. Garcia-Retamero (2017), "Cassandra's Regret: The Psychology of Not Wanting to Know," *Psychological Review*, 124 (2), 179-96.

Gittins, J. C. (1979), "Bandit Processes and Dynamic Allocation Indices," *Journal of the Royal Statistical Society: Series B (Methodological)*, 41 (2), 148-64.

Glaholt, M. G. and E. M. Reingold (2011), "Eye Movement Monitoring as a Process Tracing Methodology in Decision Making Research," *Journal of Neuroscience, Psychology, and Economics*, 4 (2), 125-46.

Glaholt, M. G., M.-C. Wu, and E. M. Reingold (2009), "Predicting Preference from Fixations," *PsychNology Journal*, 7 (2), 141-58.

--- (2010), "Evidence for Top-Down Control of Eye Movements During Visual Decision Making," *Journal of Vision*, 10 (5), 1-10.

Green, P. E. and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5 (2), 103-23.

Green, P. E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (4), 3-19.

Hanemann, W. M. (1991), "Willingness to Pay and Willingness to Accept: How Much Can They Differ?," *The American Economic Review*, 81 (3), 635-47.

Heeler, R. M., C. Okechuku, and S. Reid (1979), "Attribute Importance: Contrasting Measurements," *Journal of Marketing Research*, 16 (1), 60-63.

Hutton, S. (2008), "Cognitive Control of Saccadic Eye Movements," *Brain and Cognition*, 68 (3), 327-40.

Jaccard, J., D. Brinberg, and L. J. Ackerman (1986), "Assessing Attribute Importance: A Comparison of Six Methods," *Journal of Consumer Research*, 12 (4), 463-68.

James, W., F. Burkhardt, F. Bowers, and I. K. Skrupskelis (1890), *The Principles of Psychology*, Vol. 1: Macmillan London.

Johnson, R. M. and B. K. Orme (1996), "How Many Questions Should You Ask in Choice-Based Conjoint Studies?," *Sawtooth Software Research Paper Series*, 1-23.

Just, M. A. and P. A. Carpenter (1980), "A Theory of Reading: From Eye Fixations to Comprehension," *Psychological Review*, 87 (4), 329-54.

Ke, T. T., Z.-J. M. Shen, and J. M. Villas-Boas (2016), "Search for Information on Multiple Products," *Management Science*, 62 (12), 3576-603.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2010), "Online Demand under Limited Consumer Search," *Marketing Science*, 29 (6), 1001-23.

--- (2017), "The Probit Choice Model under Sequential Search with an Application to Online Retailing," *Management Science*, 63 (11), 3911-29.

Kim, T., K. Barasz, and L. K. John (2018), "Why Am I Seeing This Ad? The Effect of Ad Transparency on Ad Effectiveness," *Journal of Consumer Research*, 45 (5), 906-32.

Koop, G., D. J. Poirier, and J. L. Tobias (2007), *Bayesian Econometric Methods*: Cambridge University Press.

Kotowitz, Y. and F. Mathewson (1979), "Advertising, Consumer Information, and Product Quality," *The Bell Journal of Economics*, 10 (2), 566-88.

Krajbich, I., C. Armel, and A. Rangel (2010), "Visual Fixations and the Computation and Comparison of Value in Simple Choice," *Nature Neuroscience*, 13 (10), 1292-98.

Krajbich, I., D. Lu, C. Camerer, and A. Rangel (2012), "The Attentional Drift-Diffusion Model Extends to Simple Purchasing Decisions," *Frontiers in Psychology*, 3 (193), 1-18.

Krajbich, I., B. Oud, and E. Fehr (2014), "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference," *American Economic Review*, 104 (5), 501-06.

Krajbich, I. and A. Rangel (2011), "Multialternative Drift-Diffusion Model Predicts the Relationship between Visual Fixations and Choice in Value-Based Decisions," *Proceedings of the National Academy of Sciences*, 108 (33), 13852-57.

Kumar, V. and A. Man Luo (2008), "Integrating Purchase Timing, Choice, and Quantity Decisions Models," in *Review of Marketing Research*, Vol. 4, ed. K. Naresh Malhotra: Emerald Group Publishing Limited, 63-91.

LaBerge, D. (1995), *Attentional Processing: The Brain's Art of Mindfulness*, Vol. 2: Harvard University Press.

Lam, S. Y., K. Chakrabarti, G. M. Ionkov, S. H. Kim, and B. W. Witt (2012), "Recommendation System with Multiple Integrated Recommenders," *Amazon Technologies Inc Patent US8260787B2*.

Lambrecht, A. and C. Tucker (2019), "Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads," *Management Science*, 65 (7), 2966-81.

Lamkhede, S. and S. Das (2019), "Challenges in Search on Streaming Services: Netflix Case Study," *arXiv preprint arXiv:1903.04638*.

Liechty, J., R. Pieters, and M. Wedel (2003), "Global and Local Covert Visual Attention: Evidence from a Bayesian Hidden Markov Model," *Psychometrika*, 68 (4), 519-41.

Lindsen, J. P., R. Jones, S. Shimojo, and J. Bhattacharya (2010), "Neural Components Underlying Subjective Preferential Decision Making," *NeuroImage*, 50 (4), 1626-32.

Lohse, G. L. and E. J. Johnson (1996), "A Comparison of Two Process Tracing Methods for Choice Tasks," *Organizational Behavior and Human Decision Processes*, 68 (1), 28-43.

Lopez, J. S. A., M. H. Tall, B. A. G. Lefaudeux, and H. H. T. Skovsgaard (2017), "Dynamic Eye Tracking Calibration," *Facebook Patent US9851791B2*.

Luce, M. F., J. W. Payne, and J. R. Bettman (2000), "Coping with Unfavorable Attribute Values in Choice," *Organizational Behavior and Human Decision Processes*, 81 (2), 274-99.

Lynch Jr, J. G., H. Marmorstein, and M. F. Weigold (1988), "Choices from Sets Including Remembered Brands: Use of Recalled Attributes and Prior Overall Evaluations," *Journal of Consumer Research*, 15 (2), 169-84.

Maayan-Metzger, A., P. Kedem-Friedrich, and J. Kuint (2005), "To Vaccinate or Not to Vaccinate—That Is the Question: Why Are Some Mothers Opposed to Giving Their Infants Hepatitis B Vaccine?," *Vaccine*, 23 (16), 1941-48.

MacKenzie, S. B. (1986), "The Role of Attention in Mediating the Effect of Advertising on Attribute Importance," *Journal of Consumer Research*, 13 (2), 174-95.

Maćkowiak, B. A., F. Matějka, and M. Wiederholt (2018), "Survey: Rational Inattention, a Disciplined Behavioral Model," *CEPR Discussion Papers* (No. DP13243), Available at SSRN: https://ssrn.com/abstract=3266436.

Manohar, S. G. and M. Husain (2013), "Attention as Foraging for Information and Value," *Frontiers in Human Neuroscience*, 7 (1), 1-16.

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012), "Revealed Attention," *American Economic Review*, 102 (5), 2183-205.

Matějka, F. and A. McKay (2015), "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105 (1), 272-98.

McFadden, D. (1973), *Conditional Logit Analysis of Qualitative Choice Behavior*, New York: Academic Press.

Meißner, M., A. Musalem, and J. Huber (2016), "Eye Tracking Reveals Processes That Enable Conjoint Choices to Become Increasingly Efficient with Practice," *Journal of Marketing Research*, 53 (1), 1-17.

Meyer, R. J. and Y. Shi (1995), "Sequential Choice under Ambiguity: Intuitive Solutions to the Armed-Bandit Problem," *Management Science*, 41 (5), 817-34.

Motta, M., T. Callaghan, and S. Sylvester (2018), "Knowing Less but Presuming More: Dunning-Kruger Effects and the Endorsement of Anti-Vaccine Policy Attitudes," *Social Science & Medicine*, 211 (1), 274-81.

Muthén, B. (1997), "Latent Variable Modeling of Longitudinal and Multilevel Data," *Sociological Methodology*, 27 (1), 453-80.

Muthén, B., T. Asparouhov, A. Hunter, and A. Leuchter (2011), "Growth Modeling with Nonignorable Dropout: Alternative Analyses of the STAR*D Antidepressant Trial," *Psychological Methods*, 16 (1), 17-33.

Naveh, B. R. (2017), "Techniques for Emotion Detection and Content Delivery," *Facebook Patent US9681166B2*.

Orquin, J. L. and S. M. Loose (2013), "Attention and Choice: A Review on Eye Movements in Decision Making," *Acta Psychologica*, 144 (1), 190-206.

Otter, T., J. Johnson, J. Rieskamp, G. M. Allenby, J. D. Brazell, A. Diederich, J. W. Hutchinson, S. MacEachern, S. Ruan, and J. Townsend (2008), "Sequential Sampling Models of Choice: Some Recent Advances," *Marketing Letters*, 19 (1), 255-67.

Oud, B., I. Krajbich, K. Miller, J. H. Cheong, M. Botvinick, and E. Fehr (2016), "Irrational Time Allocation in Decision-Making," *Proceesing of the Royal Society B: Biological Sciences*, 283 (1822), 1-8.

Pärnamets, P., P. Johansson, L. Hall, C. Balkenius, M. J. Spivey, and D. C. Richardson (2015), "Biasing Moral Decisions by Exploiting the Dynamics of Eye Gaze," *Proceedings of the National Academy of Sciences*, 112 (13), 4170-75.

Payne, J. W. (1976), "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance*, 16 (2), 366-87.

Payne, J. W. and J. R. Bettman (2004), "Walking with the Scarecrow: The Information-Processing Approach to Decision Research," in *Blackwell Handbook of Judgment and Decision Making*, ed. Derek J. Koehler and Nigel Harvey, 110-32.

Payne, J. W., J. R. Bettman, and M. F. Luce (1996), "When Time Is Money: Decision Behavior under Opportunity-Cost Time Pressure," *Organizational Behavior and Human Decision Processes*, 66 (2), 131-52.

Pieters, R. and L. Warlop (1999), "Visual Attention During Brand Choice: The Impact of Time Pressure and Task Motivation," *International Journal of Research in Marketing*, 16 (1), 1-16.

Pieters, R. and M. Wedel (2004), "Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects," *Journal of Marketing*, 68 (2), 36-50.

--- (2007), "Goal Control of Attention to Advertising: The Yarbus Implication," *Journal of Consumer Research*, 34 (2), 224-33.

Powell, W. B. and I. O. Ryzhov (2012), *Optimal Learning*, Vol. 841: John Wiley & Sons.

R Core Team (2018), "R: A Language and Environment for Statistical Computing," R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.

Ratcliff, R. (1978), "A Theory of Memory Retrieval," *Psychological Review*, 85 (2), 59-108.

Ratcliff, R. and P. L. Smith (2004), "A Comparison of Sequential Sampling Models for Two-Choice Reaction Time," *Psychological Review*, 111 (2), 333-67.

Ratcliff, R., P. L. Smith, S. D. Brown, and G. McKoon (2016), "Diffusion Decision Model: Current Issues and History," *Trends in Cognitive Science*, 20 (4), 260-81.

Rayner, K. (1978), "Eye Movements in Reading and Information Processing," *Psychological Bulletin*, 85 (3), 618-60.

--- (1998), "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological Bulletin*, 124 (3), 372-422.

Reichle, E. D. and D. Drieghe (2015), "Using E-Z Reader to Examine the Consequences of Fixation-Location Measurement Error," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41 (1), 262-70.

Reutskaja, E., R. Nagel, C. F. Camerer, and A. Rangel (2011), "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study," *American Economic Review*, 101 (2), 900-26.

Roe, R. M., J. R. Busemeyer, and J. T. Townsend (2001), "Multialternative Decision Field Theory: A Dynamic Connectionst Model of Decision Making," *Psychological Review*, 108 (2), 370-92.

Rosbergen, E., R. Pieters, and M. Wedel (1997), "Visual Attention to Advertising: A Segment-Level Analysis," *Journal of Consumer Research*, 24 (3), 305-14.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005), *Bayesian Statistics and Marketing*: John Wiley & Sons.

Rossi, P. E., R. E. McCulloch, and G. M. Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15 (4), 321-40.

Russo, J. E. and F. Leclerc (1994), "An Eye-Fixation Analysis of Choice Processes for Consumer Nondurables," *Journal of Consumer Research*, 21 (2), 274-90.

Russo, J. E. and L. D. Rosen (1975), "An Eye Fixation Analysis of Multialternative Choice," *Memory & Cognition*, 3 (3), 267-76.

Shi, S. W., M. Wedel, and R. Pieters (2013), "Information Acquisition During Online Decision Making: A Model-Based Exploration Using Eye-Tracking Data," *Management Science*, 59 (5), 1009-26.

Shimojo, S., C. Simion, E. Shimojo, and C. Scheier (2003), "Gaze Bias Both Reflects and Influences Preference," *Nature Neuroscience*, 6 (12), 1317-22.

Shocker, A. D., M. Ben-Akiva, B. Boccara, and P. Nedungadi (1991), "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions," *Marketing Letters*, 2 (3), 181-97.

Simon, H. A. (1955), "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, 69 (1), 99-118.

--- (1971), "Designing Organizations for an Information-Rich World," in *Computers, Communications, and the Public Interest*, ed. Martin Greenberger, Baltimore: MD: The Johns Hopkins Press.

Simonson, I., J. Huber, and J. Payne (1988), "The Relationship between Prior Brand Knowledge and Information Acquisition Order," *Journal of Consumer Research*, 14 (4), 566-78.

Smith, B. and G. Linden (2017), "Two Decades of Recommender Systems at Amazon.com," *IEEE Internet Computing*, 21 (3), 12-18.

Stan Development Team (2018), "Rstan: The R Interface to Stan. R Package Version 2.18.2," http://mc-stan.org/.

Steiner, J., C. Stewart, and F. Matějka (2017), "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 85 (2), 521-53.

Stüttgen, P., P. Boatwright, and R. T. Monroe (2012), "A Satisficing Choice Model," *Marketing Science*, 31 (6), 878-99.

Sunstein, C. R. (2019), "Valuing Facebook," *Behavioural Public Policy*, 1-12.

Tobii AB (2015), "Tobii Studio User's Manual Version 3.3.2."

van der Lans, R., R. Pieters, and M. Wedel (2008a), "Eye-Movement Analysis of Search Effectiveness," *Journal of the American Statistical Association*, 103 (482), 452-61.

--- (2008b), "Research Note-Competitive Brand Salience," *Marketing Science*, 27 (5), 922-31.

van der Lans, R. and M. Wedel (2017), "Eye Movements During Search and Choice," in *Handbook of Marketing Decision Models*, ed. Berend Wierenga and Ralf van der Lans, Cham: Springer International Publishing, 331-59.

van der Lans, R., M. Wedel, and R. Pieters (2011), "Defining Eye-Fixation Sequences across Individuals and Tasks: The Binocular-Individual Threshold (BIT) Algorithm," *Behavior Research Methods*, 43 (1), 239-57.

Vehtari, A., A. Gelman, and J. Gabry (2017), "Practical Bayesian Model Evaluation Using Leave-One-out Cross-Validation and WAIC," *Statistics and Computing*, 27 (5), 1413-32.

Wang, S. S. and R. van der Lans (2018), "Modeling Gift Choice: The Effect of Uncertainty on Price Sensitivity," *Journal of Marketing Research*, 55 (4), 524-40.

Wedel, M., R. Pieters, and J. Liechty (2008), "Attention Switching During Scene Perception: How Goals Influence the Time Course of Eye Movements across Advertisements," *Journal of Experimental Psychology: Applied*, 14 (2), 129.

Wedell, D. H. and S. M. Senter (1997), "Looking and Weighting in Judgment and Choice," *Organizational Behavior and Human Decision Processes*, 70 (1), 41-64.

Weitzman, M. L. (1979), "Optimal Search for the Best Alternative," *Econometrica: Journal of the Econometric Society*, 47 (3), 641-54.

Willemsen, M. C., U. Böckenholt, and E. J. Johnson (2011), "Choice by Value Encoding and Value Construction: Processes of Loss Aversion," *Journal of Experimental Psychology: General*, 140 (3), 303-24.

Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*: MIT press.

Yang, L., O. Toubia, and M. G. de Jong (2015), "A Bounded Rationality Model of Information Search and Choice in Preference Measurement," *Journal of Marketing Research*, 52 (2), 166-83.

Zhu, S.-H. and N. H. Anderson (1991), "Self-Estimation of Weight Parameter in Multiattribute Analysis," *Organizational Behavior and Human Decision Processes*, 48 (1), 36-54.